

**ECONOMIC GROWTH CENTER  
YALE UNIVERSITY**

P.O. Box 208269  
New Haven, CT 06520-8269

<http://www.econ.yale.edu/~egcenter>



Economic Growth Center Discussion Paper No. 1056

**Networks in Conflict: Theory and Evidence from the Great War of Africa**

**Michael D. König**  
University of Zurich

**Dominic Rohner**  
University of Lausanne

**Mathias Thoenig**  
University of Lausanne

**Fabrizio Zilibotti**  
Yale University

Notes: Center discussion papers are preliminary materials circulated to stimulate discussion and critical comments.

This paper can be downloaded without charge from the Social Science Research Network Electronic Paper Collection: <https://ssrn.com/abstract=2893884>

# Networks in Conflict: Theory and Evidence from the Great War of Africa\*

Michael D. König<sup>†</sup>, Dominic Rohner<sup>‡</sup>, Mathias Thoenig<sup>§</sup>, Fabrizio Zilibotti<sup>¶</sup>

December 23, 2016

## Abstract

We study from both a theoretical and an empirical perspective how a network of military alliances and enmities affects the intensity of a conflict. The model combines elements from network theory and from the politico-economic theory of conflict. We obtain a closed-form characterization of the Nash equilibrium. Using the equilibrium conditions, we perform an empirical analysis using data on the Second Congo War, a conflict that involves many groups in a complex network of informal alliances and rivalries. The estimates of the fighting externalities are then used to infer the extent to which the conflict intensity can be reduced through (i) dismantling specific fighting groups involved in the conflict; (ii) weapon embargoes; (iii) interventions aimed at pacifying animosity among groups. Finally, with the aid of a random utility model, we study how policy shocks can induce a reshaping of the network structure.

---

\*We would like to thank Valentin Muller, Sebastian Ottinger, André Python, Timo Schaefer, Matthias Schief, Nathan Zorzi, and Laura Zwyssig for excellent research assistance. We are grateful for helpful comments to the editor, five anonymous referees, Daron Acemoglu, Jerome Adda, Alberto Alesina, Alessandra Casella, Ernesto Dal Bo, Melissa Dell, Rachel Griffith, David Hemous, Macartan Humphreys, Massimo Morelli, Benjamin Olken, María Sáez Martí, Uwe Sunde, Chris Udry, Fernando Vega Redondo, Jean-Marc Robin, David Yanagizawa-Drott, and Giulio Zanella. We owe a special thanks to Rafael Lalive for his generous help on an important coding issue discussed in the paper. We also thank conference and seminar participants at ENCORE, ESEM-Asian Meeting, IEA World Congress, INFER, LLN, NBER Summer Institute, Workshop on the Economics of Organized Crime, and at the universities of Aalto, Bocconi, Bologna, CERGE-EI, Cattolica del Sacro Cuore, Chicago, Columbia, ECARES, Ecole Polytechnique, European University Institute, Geneva, Harvard/MIT, HKU, IAE-Barcelona, IIES, IMT, INSEAD, IRES, Keio, Lausanne, LSE, Luxembourg, Manchester, Marseille-Aix, Munich, Nottingham, NYU Abu Dhabi, Oxford, Oslo, Pompeu Fabra, SAET, SciencesPo Paris, Southampton, St. Gallen, Toulouse, ULB, and USI-Lugano. Mathias Thoenig acknowledges financial support from the ERC Starting Grant GRIEVANCES-313327. Dominic Rohner gratefully acknowledges funding from the Swiss National Fund grant 100017\_150159 on “Ethnic Conflict.” Michael König and Fabrizio Zilibotti acknowledge financial support from the Swiss National Science Foundation through the research grant 100018\_140266, and Michael König further acknowledges financial support from the Swiss National Science Foundation grant PZ00P1\_154957 / 1.

<sup>†</sup>Department of Economics, University of Zurich. Email: michael.koenig@econ.uzh.ch.

<sup>‡</sup>Department of Economics, University of Lausanne. Email: dominic.rohner@unil.ch.

<sup>§</sup>Department of Economics, University of Lausanne. Email: mathias.thoenig@unil.ch.

<sup>¶</sup>Department of Economics, Yale University. Email: zilibotti@gmail.com.

## 1 INTRODUCTION

Alliances and enmities among armed actors – be they rooted in history or in mere tactical considerations – are part and parcel of warfare. In many episodes, especially in civil conflicts, they are shallow links that are not sanctioned by formal treaties or war declarations. Even allied groups retain separate agendas and pursue self-interested goals in competition with each other. Understanding the role of informal networks of military alliances and enmities is important, not only for predicting outcomes, but also for designing and implementing policies to contain or put an end to violence. Yet, with only few exceptions, the existing political and economic theories of conflict restrict attention to a small number of players, and do not consider network aspects. In this paper, we construct a stylized theory of conflict that captures the effect of informal networks of alliances and enmities, and apply it to the empirical study of the Second Congo War and its aftermath.

The theoretical benchmark is a *contest success function*, henceforth CSF, in the vein of Tullock (1980). In a standard CSF, the share of the prize accruing to a group is determined by the effort (*fighting effort*) that this group commits to the conflict relative to its contenders. In our model, the network of alliances and enmities modifies the sharing rule of the CSF by introducing novel externalities. More precisely, we assume that the share of the prize accruing to group  $i$  is determined by the group’s relative strength, which we label *operational performance*. In turn, this is determined by group  $i$ ’s own fighting effort and by the fighting effort of its allied and enemy groups. The fighting effort of group  $i$ ’s allies increases group  $i$ ’s operational performance, whereas the fighting effort of its enemies decreases it. Each group decides its effort non-cooperatively. Since the cost of fighting is borne individually by each group, a motive for strategic behavior arises among both enemies and allies. The complex externality web affects the optimal fighting effort of all groups.

We provide an analytical solution for the Nash equilibrium of the game. The model can be used to predict how the network structure of alliances and rivalries affects the overall conflict intensity, given by the sum of the fighting efforts of all contenders, which is our measure of the welfare loss. Network externalities are a driver of the escalation or containment of violence.

The empirical analysis, which is based on the structural equations of the model, focuses on the Second Congo War, a large-scale conflict involving a rich network of informal alliances and rivalries that started in 1998 in the Democratic Republic of Congo (DRC). To identify the network, we use information from a variety of expert and data sources. The estimated network features numerous intransitivities, showing that this conflict cannot be described as the clash between two unitary camps (see Figure 2 below).

Our estimation strategy exploits panel variations in the yearly number of clashes involving 80 armed groups in 1998–2010, as measured by the Armed Conflict Location Event Database (ACLED). A group’s fighting effort is proxied by the number of clashes in which it is involved. Controlling for group fixed effects, we regress each group’s fighting effort on the total fighting efforts of its degree-one allies and enemies, respectively. Since these efforts are endogenous and subject to a reflection problem, we adopt an instrumental variable (IV) strategy similar to that used by Acemoglu *et al.* (2015). Our identification strategy exploits the exogenous variation in the average weather conditions facing, respectively, the set of allies and of enemies of each group. The focus on weather shocks is motivated by the recent literature documenting that these have important effects on fighting intensity (see Dell 2012, Hidalgo *et al.* 2010, Hsiang *et al.* 2013, Miguel *et al.* 2004, and Vanden Eynde 2017). Without imposing any restriction, we find that the two estimated externalities have the opposite sign pattern which aligns with the predictions of the theory. Moreover, we find no external effect from neutral groups, which is also in line with the theoretical predictions.

After estimating the network externalities, we perform a variety of counterfactual policy experiments. First, we consider targeted policies that either induce some groups to drop out of the conflict or increase their marginal cost of fighting (e.g., arms embargoes). The analysis singles out armed groups whose decommissioning or weakening is most effective for scaling down conflict. Second, we study the effect of pacification policies aimed at reducing the hostility between enemy groups, e.g., through bringing selected actors to the negotiating table. Since enmities tend to increase the conflict intensity, bilateral or multilateral pacifications tend to reduce violence. We find that the gains from pacification policies can be large. At instances, the reduction in the level of armed activity is well in excess of the amount of fighting between the groups whose bilateral hostilities were placated.

The results highlight the key role of Rwanda and Uganda in the conflict, although some smaller guerrilla groups such as the Lord Resistance Army (LRA) are also important drivers of violence. Arms embargoes that increase the fighting cost of groups without inducing them to demobilize are generally ineffective because the reduction in the targeted groups' activity is typically offset by an increase in the activity of the other groups. In contrast, targeted bilateral or multilateral pacification policies can be highly effective.

In most of the paper, we maintain the assumption of an exogenous network. This assumption is relaxed in an extension, where we allow the network to adjust endogenously to policy shocks, based on the predictions of a random utility model. The recomposition of the network magnifies the effect of interventions targeting foreign groups. Removing all foreign groups reduces the conflict by 41%, significantly more than in the case of an exogenous network (27%). These results are in line with the narrative that foreign intervention is an important driver of the DRC conflict.

Our contribution is related to various strands of the existing literature. It is linked to the growing literature on the economics of networks (e.g., Bramoullé *et al.* 2014, Jackson 2008). Franke and Öztürk (2015) and Huremovic (2014) study strategic interactions of multiple agents in conflict networks. Two recent papers by Hiller (2017) and Jackson and Nei (2014) study the endogenous formation of networks in conflict models. None of these papers endogenizes the choice of fighting effort. Also, while these articles are theoretical, our study provides a quantification of the theory by estimating the key network externalities based on the structural equations of the theory. The empirical strategy is related to Acemoglu *et al.* (2015), who estimate a political economy model of public goods provision using a network of Colombian municipalities. Further, one of our policy experiments is an application of the key player analysis in Ballester *et al.* (2006).

Our study is broadly related to the growing politico-economic literature on conflict. The papers in this literature typically focus on two groups confronting each other (see, e.g., Rohner *et al.* 2013). A number of studies use a CSF (see, e.g., Grossman and Kim 1995, Hirshleifer 1989, and Skaperdas 1996), while Esteban and Ray (2001) consider collective action problems across multiple groups. Olson and Zeckhauser (1966) consider free riding problems in alliances. Bates *et al.* (2002) distinguish between fighting and arming, which is related to our analysis of the effects of arms embargoes. Bloch (2012) and Konrad (2011) provide excellent surveys of the literature.

Finally, our paper is related to the empirical literature on civil war, and in particular to the recent literature that studies conflict using very disaggregated micro-data on geolocalized fighting events, such as for example Cassar *et al.* (2013), Dube and Vargas (2013), La Ferrara and Harari (2012), Michalopoulos and Papaioannou (2013), Rohner *et al.* (2013b), and (specifically on the DRC conflict) Sanchez de la Sierra (2016).

The paper is organized as follows: Section 2 presents the theory; Section 3 discusses the context of the Second Congo War and the data; Section 4 presents the estimation results. Section 5 performs policy counterfactual analyses; Section 6 proposes an extension where the network structure

responds to policy shocks. Appendixes A–B contain technical and empirical details, respectively. Appendix C (available from the authors’ webpages) contains additional material.

## 2 THEORY

### 2.1 ENVIRONMENT

We consider a population of  $n \in \mathbb{N}$  agents (henceforth, *groups*) whose interactions are captured by a network  $G \in \mathcal{G}^n$ , where  $\mathcal{G}^n$  denotes the class of graphs on  $n$  nodes. Each pair of groups can be in one of three states: alliance, enmity, or neutrality. We represent the set of bilateral states by the signed adjacency matrix  $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq n}$  associated with the network  $G$ , where, for all  $i \neq j$ ,

$$a_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are allies,} \\ -1, & \text{if } i \text{ and } j \text{ are enemies,} \\ 0, & \text{if } i \text{ and } j \text{ are in a neutral relationship.} \end{cases}$$

We conventionally set  $a_{ii} = 0$  and define the number of group  $i$ ’s allies and enemies as  $d_i^+ \equiv \sum_{j=1}^n a_{ij}^+$  and  $d_i^- \equiv \sum_{j=1}^n a_{ij}^-$ , respectively, where  $a_{ij}^+ = 1$  if  $a_{ij} = 1$ , and  $a_{ij}^+ = 0$  otherwise, and similarly  $a_{ij}^- = 1$  if  $a_{ij} = -1$ , and  $a_{ij}^- = 0$  otherwise.

The  $n$  groups compete for a divisible prize denoted by  $V$ , which can be interpreted as the control over territories and natural resources. We assume payoffs to be determined by a generalized Tullock CSF, which maps the groups’ relative fighting intensities into shares of the prize. More formally, we postulate a payoff function  $\pi_i : \mathcal{G}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  given by

$$\pi_i(G, \mathbf{x}) = \begin{cases} \frac{\varphi_i(G, \mathbf{x})}{\sum_{j=1}^n \max\{0, \varphi_j(G, \mathbf{x})\}} V - x_i, & \text{if } \varphi_i(G, \mathbf{x}) \geq 0, \\ -D, & \text{if } \varphi_i(G, \mathbf{x}) < 0. \end{cases} \quad (1)$$

The vector  $\mathbf{x} \in \mathbb{R}^n$  describes the fighting effort of each group (the choice variable), whereas  $\varphi_i \in \mathbb{R}$  is group  $i$ ’s *operational performance* (OP). The parameter  $D \geq 0$  is the *defeat cost* that groups suffer when their OP falls below zero (as discussed below). Group  $i$ ’s OP is assumed to depend on group  $i$ ’s fighting effort  $x_i$ , as well as on its allies’ and enemies’ efforts. More formally, we assume that

$$\varphi_i(G, \mathbf{x}) = x_i + \beta \sum_{j=1}^n (1 - \mathbb{1}_{\mathcal{D}}(j)) a_{ij}^+ x_j - \gamma \sum_{j=1}^n (1 - \mathbb{1}_{\mathcal{D}}(j)) a_{ij}^- x_j, \quad (2)$$

where  $\beta, \gamma \in [0, 1]$  are linear spillover effects from allies’ and enemies’ fighting efforts, respectively.  $\mathbb{1}_{\mathcal{D}}(j) \in \{0, 1\}$  is an indicator function that takes the unit value for groups accepting defeat and paying the cost  $D$  – these groups are assumed to exert no externality. For simplicity, with a slight abuse of notation, we henceforth set  $x_j = 0$  in equation (2) when group  $j$  accepts defeat, and omit the indicator function. The assumption that the spillovers enter linearly into the expression of  $\varphi_i$  allows us to obtain a closed-form solution. It entails some important limitations insofar as it rules out complementarities (e.g., in the effort of different allied groups) that can be important in reality.

Equation (2) postulates that each active group’s OP is an increasing (decreasing) function of the total effort exerted by its allies (enemies). These externalities compound with the one embedded in the standard CSF, which equation (2) nests as a particular case when  $a_{ij}^+ = a_{ij}^- = 0$  for all  $i$  and

$j$ .<sup>1</sup> In the rest of the paper, we normalize  $V$  to unity.<sup>2</sup>

Consider, finally, the defeat option. When the OP turns negative (for instance, because the enemies exert high effort), a group waves the white flag and suffers the defeat cost  $D$ .<sup>3</sup> This is a natural assumption: too low an OP exposes groups to other armed groups' looting and ransacking.

## 2.2 NASH EQUILIBRIUM

Each group chooses effort,  $x_i$ , *non-cooperatively* so as to maximize  $\pi_i(G, [x_i, \mathbf{x}_{-i}])$ , given  $\mathbf{x}_{-i}$ . The Nash equilibrium is a fixed point of the effort vector.

Consider a candidate equilibrium where  $\hat{n} \leq n$  groups participate actively in the contest. A necessary and sufficient condition for the optimal effort choice to be a concave problem is that, for  $i = 1, 2, \dots, \hat{n}$ ,

$$\frac{\partial}{\partial x_i} \sum_{j=1}^{\hat{n}} \varphi_j = 1 + \beta d_i^+ - \gamma d_i^- > 0. \quad (3)$$

In the empirical analysis below, we check that this condition holds in the empirical network for our estimates of  $\beta$  and  $\gamma$ . When condition (3) holds, the optimal effort choice of participants satisfies a system of First-Order Conditions (FOCs). Using equations (1)-(2), one obtains:

$$\frac{\partial \pi_i(G, \mathbf{x})}{\partial x_i} = 0 \iff \varphi_i = \frac{1}{1 + \beta d_i^+ - \gamma d_i^-} \left( 1 - \sum_{j=1}^{\hat{n}} \varphi_j \right) \sum_{j=1}^{\hat{n}} \varphi_j.$$

Rearranging terms allows us to obtain a simple expression for the equilibrium OP level

$$\varphi_i^*(G) = \Lambda^{\beta, \gamma}(G) \left( 1 - \Lambda^{\beta, \gamma}(G) \right) \Gamma_i^{\beta, \gamma}(G), \quad (4)$$

and for the equilibrium share of the prize,

$$\frac{\varphi_i^*(G)}{\sum_{j=1}^{\hat{n}} \varphi_j^*(G)} = \frac{\Gamma_i^{\beta, \gamma}(G)}{\sum_{j=1}^{\hat{n}} \Gamma_j^{\beta, \gamma}(G)}, \quad (5)$$

where

$$\Gamma_i^{\beta, \gamma}(G) \equiv \frac{1}{1 + \beta d_i^+ - \gamma d_i^-} > 0 \text{ and } \Lambda^{\beta, \gamma}(G) \equiv 1 - \frac{1}{\sum_{i=1}^{\hat{n}} \Gamma_i^{\beta, \gamma}(G)}. \quad (6)$$

The term  $\Gamma_i^{\beta, \gamma}(G) > 0$  is a measure of the local hostility level capturing the externalities associated with group  $i$ 's first-degree alliance and enmity links. One can show that  $0 < \Lambda^{\beta, \gamma}(G) < 1$ , implying that  $\varphi_i^*(G) > 0$ . Moreover, both  $\Gamma_i^{\beta, \gamma}(G)$  and  $\Lambda^{\beta, \gamma}(G)$  are decreasing with  $\beta$  and increasing with  $\gamma$ . Equation (5) implies that the share of the prize accruing to group  $i$  increases in the number of its allies and decreases in the number of its enemies.

The next proposition (proof in Appendix A.1) characterizes the equilibrium.

---

<sup>1</sup>Consider, for example, an alliance link,  $a_{kk'} = 1$ . Then, an increase in the effort of  $k'$  affects the payoff of  $k$  via two channels: (i) the standard negative externality working through the denominator; (ii) the positive externality working through the numerator. Thus, holding efforts constant, an alliance between two groups increases the share of the prize jointly accruing to them, at the expense of the remaining groups. To the opposite, enmity links strengthen the negative externality of the standard CSF.

<sup>2</sup>Note that, in equilibrium, both  $x_i$  and  $\pi_i$  are proportional to  $V$ .

<sup>3</sup>Note that we do not impose any non-negativity constraint on  $x_i$ . Given the linearity of the payoff function, the zero effort level is a matter of normalization.

**Proposition 1.** Let  $\mathbf{A}^+ = (a_{ij}^+)_{1 \leq i, j \leq n}$  and  $\mathbf{A}^- = (a_{ij}^-)_{1 \leq i, j \leq n}$ , implying that  $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$ , and denote the corresponding subgraphs as  $G^+$  and  $G^-$ , respectively, so that  $G = G^+ \oplus G^-$ . Assume that  $\beta + \gamma < 1 / \max\{\lambda_{\max}(G^+), d_{\max}^-\}$ , where  $\lambda_{\max}(G)$  denotes the largest eigenvalue associated with the adjacency matrix of the network  $G$ , and that condition (3) holds true for all  $i = 1, 2, \dots, n$ . Then,  $\exists \underline{D} < \infty$  such that,  $\forall D > \underline{D}$ , there exists an interior Nash equilibrium such that,  $\forall i = 1, 2, \dots, n$ , the equilibrium effort levels and OPs are given by

$$x_i^*(G) = \Lambda^{\beta, \gamma}(G) \left(1 - \Lambda^{\beta, \gamma}(G)\right) c_i^{\beta, \gamma}(G), \quad (7)$$

and  $\varphi_i = \varphi_i^*(G) \geq 0$  as given by equation (4), for  $\hat{n} = n$ . Here,  $\Gamma_i^{\beta, \gamma}(G)$  and  $\Lambda^{\beta, \gamma}(G)$  are defined by equation (6), and

$$\mathbf{c}^{\beta, \gamma}(G) \equiv (\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-)^{-1} \mathbf{\Gamma}^{\beta, \gamma}(G) \quad (8)$$

is a centrality vector whose generic element  $c_i^{\beta, \gamma}(G)$  describes group  $i$ 's centrality in the network  $G$ . Finally, the equilibrium payoffs are given by

$$\pi_i^*(G) = (1 - \Lambda^{\beta, \gamma}(G)) \left(\Gamma_i^{\beta, \gamma}(G) - \Lambda^{\beta, \gamma}(G) c_i^{\beta, \gamma}(G)\right) > -D. \quad (9)$$

If, in addition,  $\Gamma_i^{0, \gamma} > 0$  for all  $i = 1, \dots, n$ , then,  $\exists \underline{\underline{D}}$  (where  $\underline{D} \leq \underline{\underline{D}} < \infty$ ) such that,  $\forall D > \underline{\underline{D}}$ , the equilibrium is unique.

The first part of the proposition yields an existence result. Condition (i) is a sufficient condition for the matrix in (8) to be invertible. Equation (7) follows from the set of FOCs. For a sufficiently large  $D$ , an equilibrium exists where all groups participate in the contest.<sup>4</sup> Figure A.1 in the appendix provides an illustration of the existence result by displaying the payoff function  $\pi_i(G, [x_i, \mathbf{x}_{-i}^*])$  at the equilibrium strategy profile.

The second part of the proposition establishes that, under a stronger set of conditions, the Nash equilibrium where all agents participate is unique. In this case, setting  $D$  sufficiently high rules out equilibrium configurations in which a partition of groups takes the defeat option. For lower values of  $D$ , equilibria in which some groups accept defeat may instead exist, and multiple equilibria are possible.

### 2.3 CENTRALITY AND WELFARE

The centrality measure  $c_i^{\beta, \gamma}(G)$  plays a key role in Proposition 1. In particular, the ratio between the fighting efforts of any two groups equals the ratio between the respective centralities in the network:

$$\frac{x_i^*(G)}{x_j^*(G)} = \frac{c_i^{\beta, \gamma}(G)}{c_j^{\beta, \gamma}(G)}.$$

In Appendix A.2, we relate our centrality measure to the Katz-Bonacich centrality. Moreover, we show that for networks in which the spillover parameters  $\beta$  and  $\gamma$  are small, each group's fighting effort increases, and each group's equilibrium payoff decreases in the weighted difference between the number of enmities (weighted by  $\gamma$ ) and of alliances (weighted by  $\beta$ ), i.e.,  $d_i^- \gamma - d_i^+ \beta$ . Intuitively, a group with many enemies tends to fight harder and to appropriate a smaller share of the prize, whereas a group with many allies tends to fight less and to appropriate a large share of the prize.

---

<sup>4</sup>The focus on an equilibrium in which all groups are active is without loss of generality. The results are identical if there are  $\tilde{n} \leq n$  groups waving the white flag. In our model, inactive groups exert no externality, and can therefore be ignored.

To discuss normative implications of the theory, we define the total *rent dissipation* associated with the conflict:

$$\text{RD}^{\beta,\gamma}(G) \equiv \sum_{i=1}^n x_i^*(G) = \Lambda^{\beta,\gamma}(G)(1 - \Lambda^{\beta,\gamma}(G)) \sum_{i=1}^n c_i^{\beta,\gamma}(G). \quad (10)$$

Clearly, minimizing rent dissipation is equivalent to maximizing welfare.

In Appendix C, we also provide a simple illustration of the role of alliances and enmities for conflict escalation in the particular case of *regular* networks. *Regular* networks have the property that every group  $i$  has  $d_i^+ = k^+$  alliances and  $d_i^- = k^-$  enmities, namely, all groups have the same centrality. In this case, one can show easily that effort and rent dissipation decrease (increase) in the number of alliances (enmities).

## 2.4 HETEROGENEOUS FIGHTING TECHNOLOGIES

So far, we have maintained that all groups have access to the same OP technology. This was useful for focusing sharply on the network structure. In reality, armed groups typically differ in size, wealth, access to arms, leadership, etc. In this section, we generalize our model by allowing fighting technologies to differ across groups. We restrict attention to additive heterogeneity, since this is crucial for achieving identification in the econometric model presented below. Suppose that group  $i$ 's OP is given by:

$$\varphi_i(G, \mathbf{x}) = \tilde{\varphi}_i + x_i + \beta \sum_{j=1}^n a_{ij}^+ x_j - \gamma \sum_{j=1}^n a_{ij}^- x_j, \quad (11)$$

where  $\tilde{\varphi}_i$  is a group-specific shifter affecting the OP. Equation (11) will be the basis of our econometric analysis, where we introduce both observable and unobservable sources of heterogeneity.

In Appendix A.3, we show that the equilibrium OP is unchanged, and continues to be given by equation (4). Likewise, equation (5) continues to characterize the share of the prize appropriated by each group. Somewhat surprising, the share of resources appropriated by group  $i$ ,  $\varphi_i^*(G) / \sum_{j=1}^n \varphi_j^*(G)$ , is independent of  $\tilde{\varphi}_i$ . However,  $\tilde{\varphi}_i$  affects the equilibrium effort exerted by group  $i$  and its payoff. In particular, the vector of the equilibrium fighting efforts is now given by

$$\mathbf{x}^* = (\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-)^{-1} (\Lambda^{\beta,\gamma}(G)(1 - \Lambda^{\beta,\gamma}(G)) \mathbf{\Gamma}^{\beta,\gamma}(G) - \tilde{\boldsymbol{\varphi}}), \quad (12)$$

where  $\tilde{\boldsymbol{\varphi}} = (\tilde{\varphi}_i)_{1 \leq i \leq n}$  is the vector of group-specific shifters, and the definitions of  $\Lambda^{\beta,\gamma}(G)$  and  $\mathbf{\Gamma}^{\beta,\gamma}(G)$  are unchanged (see Proposition 2).

## 3 EMPIRICAL APPLICATION - THE SECOND CONGO WAR

In the rest of the paper, we focus on the recent civil conflict in the DRC with the goal of providing a quantitative evaluation of the theory. More specifically, we estimate the externality parameters  $\beta$  and  $\gamma$  from the structural equation (11) characterizing the Nash equilibrium of the model. Then, we perform counterfactual policy experiments and assess their effectiveness in scaling down conflict. We start by discussing the historical context of the Second Congo War, the data sources, and the network structure.

### 3.1 HISTORICAL CONTEXT

The DRC is the largest Sub-Saharan African country in terms of area, and is populated by about 75 million inhabitants. It is a failed state. After gaining independence from Belgium in 1960, it



experienced recurrent political instability and wars that turned it into one of the poorest countries in the world, in spite of the abundance of natural resources. The DRC is an ethnically fragmented country with over 200 ethnic groups. The Congo conflict involves many inter-connected domestic and foreign actors: three Congolese rebel movements, 14 foreign armed groups, and a countless number of militias (Autesserre 2008).

The war in Congo is intertwined with the ethnic conflicts in neighboring Rwanda and Uganda. In 1994, Hutu radicals took control of the Rwandan government and allowed ethnic militias to perpetrate the mass killing of nearly a million Tutsis and moderate Hutus in less than one hundred days. After losing power to the Tutsi-led Rwandan Patriotic Front, over a million Hutus fled Rwanda and found refuge in the DRC, ruled at that time by the dictator Mobutu Sese Seko. The refugee camps hosted, along with civilians, former militiamen and genocidaires who clashed regularly with the local Tutsi population, most notably in the Kivu region (Seybolt 2000).

As ethnic tensions mounted, a large coalition of African countries centered on Uganda and Rwanda supported an anti-Mobutu rebellion led by Laurent-Désiré Kabila. The First Congo War (1996-97) ended with Kabila's victory. However, his relationship with his former sponsors soon turned sour. Resenting the enormous political influence exerted by the two neighbors, Kabila first dismissed his Rwandan chief of staff, James Kabarebe, and then ordered all Rwandan and Ugandan troops to leave the country. New ethnic clashes soon erupted in Eastern Congo. The crisis escalated into outright war. Rwanda and Uganda mobilized the local Tutsi population and armed a well-organized rebel group, the Rally for Congolese Democracy (RCD), that took control of Eastern Congo. The main Hutu military organization, the Democratic Forces for the Liberation of Rwanda (FDLR), sided with Kabila, who also received the international support of Angola, Chad, Namibia, Sudan, and Zimbabwe, and of the local Mayi-Mayi militias.

Officially, the Second Congo War ended in July 2003. In reality, stable peace was never achieved (Stearns 2011), and fighting is still going on today. The conflict is highly fragmented. In Prunier's words, "*the continent was fractured, not only for or against Kabila, but within each of the two camps*" (Prunier 2011: 187). Similarly, there was in-fighting among different pro-government paramilitary groups, such as the Mayi-Mayi militias. The DRC army (FARDC) themselves were notoriously prone to internal fights and mutinies, spurred by the fact that its units are segregated along ethnic lines and often correspond to former ethnic militias or paramilitaries that got integrated into the national army. In summary, far from being a war between two unitary camps, the conflict engaged a complex web of informal alliances and enmities with many non-transitive links.

After a major reshuffling at the end of the First Congo War, the web of alliances and enmities between the main armies and rebel groups has remained fairly stable in the period 1998-2010 (see Prunier 2011: 187ff). Yet, there were some notable exceptions. The relationship between Uganda and Rwanda cracked soon after the start of the conflict, culminating in a series of armed confrontations in the Kisangani area in the second half of 1999 and in 2000 that caused the death of over 600 civilians (see Turner 2007: 200). The crisis spilled over to the local proxy of the two countries: the RCD split into the Uganda-backed RCD-Kisangani (RCD-K) and the Rwanda-backed RCD-Goma (RCD-G). After 2000, the relationship between Uganda and Rwanda lived in a knife-edge equilibrium where recurrent tensions and skirmishes were prevented from spiraling into a full-scale conflict (McKnight 2015).

The relationship between the FARDC and the FDLR is also troubled. In the earlier stage of the conflict, they were allies. Things changed after Laurent Kabila was assassinated in 2001. In 2002, a peace agreement signed at Sun City in South Africa allowed Joseph Kabila, Laurent's son, to remain in power in exchange for his commitment to end the support for anti-Rwandan rebel armies in Congo. As a result, the relations with the FARDC became volatile. The FDLR

kept engaging Tutsi forces in the Kivu region, raising concern for a new full-fledged intervention of Rwanda. In 2009, the FDLR attacked civilians in some Kivu villages, prompting a major joint military operation of the FARDC and Rwanda against them.<sup>5</sup>

Finally, there were numerous local rebellions which led to the formation of new groups and break-away mutinies of pre-existing militias. An example is the 2002 revolt of the Banyamulenge, a Congolese Tutsi group, originating from a mutiny led by Patrick Masunzu that the mainstream Rwanda-backed RCD-G troops failed to crush.

### 3.2 DATA

We use a panel of annual observations for the period 1998–2010 drawing on a variety of data sources.<sup>6</sup> The unit of analysis is at the group×year level. The summary statistics are displayed in Table B.1. In the rest of this section, we provide a summary description of the dataset. More details about the data construction can be found in Appendix C.

**Groups** – The main data source for the fighting effort and geolocalization of groups is ACLED ([www.acleddata.com](http://www.acleddata.com)).<sup>7</sup> This dataset contains 4676 geolocalized violent events taking place in the DRC involving on the whole 80 groups: 4 Congolese state army groups, 47 domestic Congolese non-state militias, 11 foreign government armies, and 18 foreign non-state militias.<sup>8</sup> A complete list of the groups is provided in Appendix C.

Our classification of groups strictly follows ACLED. The cases of the FARDC and of Rwanda deserve a special mention. ACLED codes each of these two actors as split into two groups by the period of activity: “FARDC (1997-2001) (Kabila, L.)” (henceforth, FARDC-LK) and “FARDC (2001-) (Kabila, J.)” (henceforth, FARDC-JK); “Rwanda (1994-1999)” (henceforth, Rwanda-I) and “Rwanda (1999-)” (henceforth, Rwanda-II). In the case of the FARDC, the split is determined by the assassination of Laurent Kabila, followed by the peace agreement of 2002 that marks the official end of the Second Congo War. In the case of Rwanda, the threshold coincides with the deterioration of the relationships between Uganda and Rwanda. In our baseline estimation, we do not merge these groups since the discontinuities reflect genuine political breaking points. However, we also consider robustness checks in which these groups are merged.

**Fighting events** – For each event, ACLED provides information on the location, date, and identity of the groups involved on the same and on opposite sides. ACLED draws primarily on three types of sources: information from local, regional, national, and continental media, reviewed on a daily basis; NGO reports; Africa-focused news reports and analyses. ACLED data is subject to measurement error in two dimensions. First, many events go unreported (see Van der Windt and Humphreys 2016, discussed in more detail in Appendix C). Second, the precision of the geolocalization of ACLED events has been questioned (Eck 2012).<sup>9</sup> To address the geolocalization issue, we supplement the information in ACLED with the UCDP Georeferenced Event Dataset

---

<sup>5</sup>See BBC News, 13 May 2009, <http://news.bbc.co.uk/2/hi/africa/8049105.stm>.

<sup>6</sup>The main reason why our sample ends in 2010 is that, after this date, the intensity of conflict decreases significantly (in 2010 there were 301 events, while in 2011 only 89), and there was some reshuffling of alliances. In addition, the version of the rainfall data we use is only available until 2010.

<sup>7</sup>Recent papers using ACLED data include, among others, Berman *et al.* (2014), Cassar *et al.* (2013), Michalopoulos and Papaioannou (2013), and Rohner *et al.* (2013b).

<sup>8</sup>We only include organized armed groups in the dataset and exclude other actors such as civilians. While we start out with 100 fighting groups in the raw data, we drop all non-bilateral fighting events where no armed group is involved in one of the two camps (e.g., events where an armed group attacks civilians). This leaves us with 80 fighting groups in the final sample.

<sup>9</sup>In a case study of Algeria 1997, she documents that 30% of events contain inaccurate geolocalization information,

(GED) (Sundberg and Melander 2013). This dataset contains detailed georeferenced information on conflict events, but contrary to ACLED the GED dataset only reports one armed group involved for each side of the conflict. In addition, there are much fewer events (i.e., ca. one third) in GED than in ACLED.

Our main dependent variable is group  $i$ 's yearly *Fighting Effort*,  $x_{it}$ . This is measured as the sum over all ACLED fighting events involving group  $i$  in year  $t$ . In the robustness section, we construct alternative fighting effort measures by restricting the count to the more conspicuous events such as those classified by ACLED as battles or those involving fatalities.

**Rainfall** – For the purpose of our IV strategy, we build the yearly average of rainfall in each group's homeland. We use a gauge-based rainfall measure from the Global Precipitation Climatology Centre (GPCC, version 6, see Schneider *et al.* 2011), at a spatial resolution of  $0.5^\circ \times 0.5^\circ$  grid-cells. A group's homeland corresponds to the spatial zone of its military operations (i.e., the convex hull containing all geolocalized ACLED events involving that group at any time during the period 1998-2010). Then, for each year  $t$ , we compute the average rainfall in the grid-cell of the homeland centroid. Figure 1 displays the fighting intensity and average climate conditions for different ethnic homelands in the DRC. Weather conditions vary considerably both across regions and over time.

In the robustness section, we also consider satellite-based rainfall measures. These use atmospheric parameters (e.g., cloud coverage, light intensity) as indirect measures of rainfall, blended with some information from local gauges. The first comes from the Global Precipitation Climatology Project (GPCP) from NOAA. The second dataset is the Tropical Rainfall Measuring Mission (TRMM) from NASA.

**Covariates** – Given the large number of groups and years, we can control for time-varying shocks affecting groups with common characteristics. To this aim, we interact time dummies with three group-specific dummies. The first is a dummy for *Government Organization* that switches on for groups that are officially affiliated to a domestic or foreign government. This dummy covers 15 groups. The second is a dummy labelled *Foreign* that switches on for 29 groups which are coded as foreign actors. The third is a dummy labelled *Large* that switches on for groups that have at least 10 enmity links (this corresponds to the 90th percentile of groups with non-zero numbers of enemies). Since some results are sensitive to the definition of this variable, we also show results with the alternative variable *Large\_6* that switches on for groups that have at least 6 enmity links (this corresponds to the 85th percentile of groups with non-zero numbers of enemies).

### 3.3 THE NETWORK

Our primary sources of information to infer the network of enmities and alliances are: (i) the Yearbook of the Stockholm International Peace Research Institute, SIPRI (Seybolt 2000), (ii) "Non-State Actor Data" (Cunningham *et al.* 2013), (iii) Briefing on the Congo War by the International Crisis Group (1998), and (iv) Williams (2013). The four sources are mutually consistent and complementary. Groups are classified as allies or enemies not only on the basis of ground fighting, but also on that of political and logistical support (in particular, they include actors operating in different parts of the country). The four expert sources allow us to code 80 alliances and 24

---

6% of events are double-counted, and 2% missing. However, the precision is higher in a case study of Burundi 2000. There, she finds 9% of observations to have inaccurate longitude/latitude information, and 2% of observations to be double-counted, with no events missing.

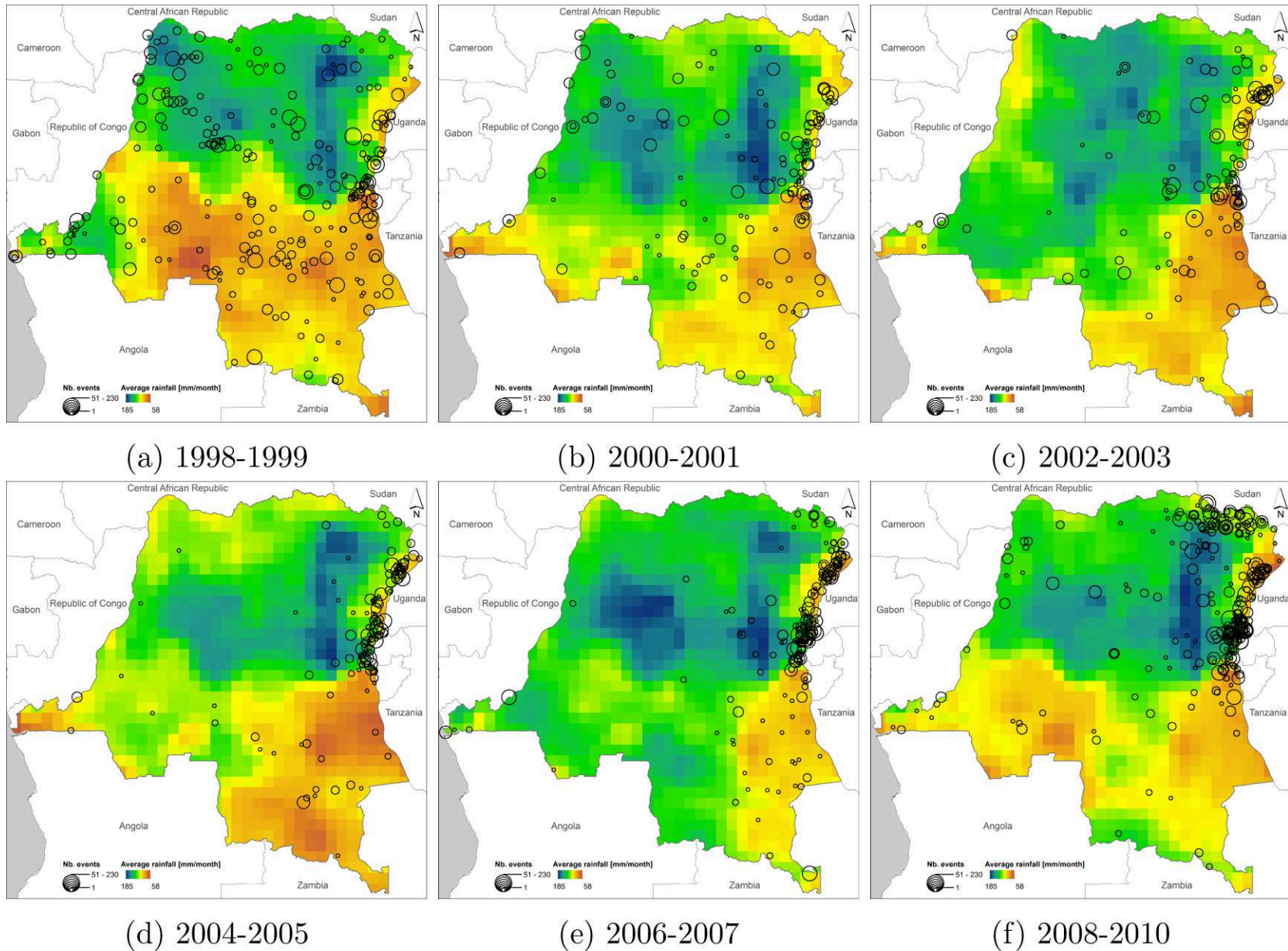


Figure 1: Map of the DRC with average rainfall data from GPCP at a spatial resolution of  $0.5^\circ \times 0.5^\circ$  grid-cells, and number of violent events from ACLED. Warmer colors (i.e., red) mean less rain, colder colors (i.e., blue) mean more rain. Larger circles mean more fighting events.

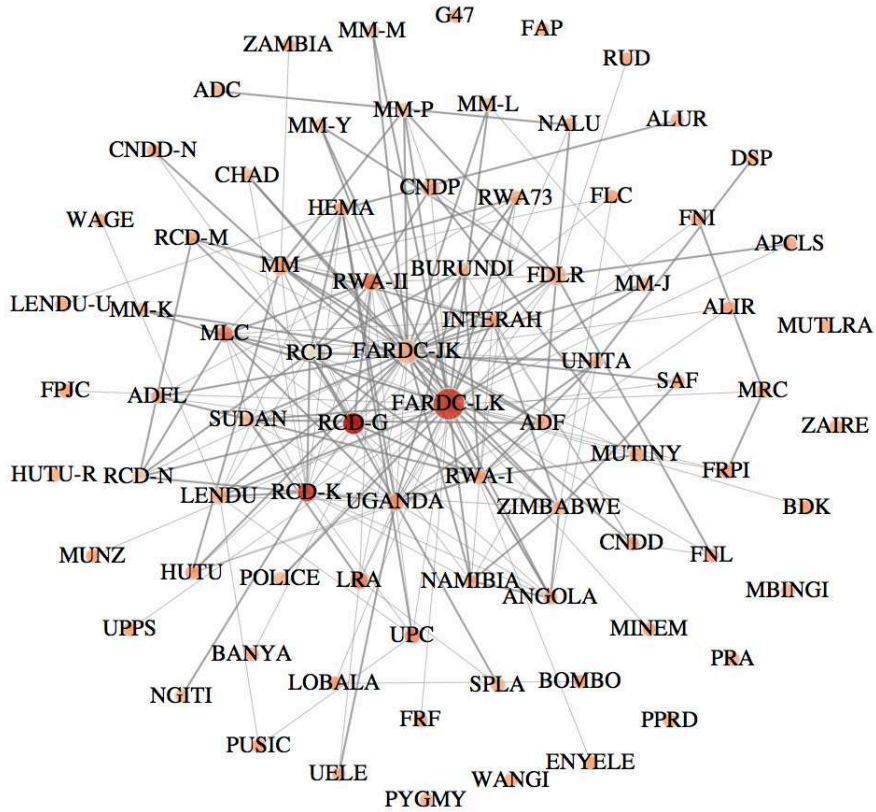


Figure 2: The figure displays the network of alliances and enmities between 80 fighting groups active in the DRC over the 1998-2010 period.

enmities.<sup>10</sup>

The main limitation of the expert coding is that it does not cover small armed groups and militias. For this reason, we complement its information with that inferred from the battlefield behavior in ACLED – where ACLED is strictly subordinate to the four expert sources. In particular, we consider two groups ( $i, j$ ) as *enemies* if they have been observed fighting on opposite sides on at least two occasions, and they have never been observed fighting as brothers in arms. Conversely, we code two groups as *allies* if they have been observed fighting on the same side in at least one occasion during the sample period, and if, in addition, they do not reach the enmity threshold of two clashes against each other. We code all other dyads as *neutral*. A concern might be that the construction of the network relies in part on the same ACLED data that we use to measure the outcome variable. In this respect, one must emphasize two important features. First, for constructing the network we exploit the bilateral information which is not used to construct the outcome variable. Second, the network is assumed to be time-invariant (at least in the baseline specification), whereas our econometric identification exploits the time variations in fighting efforts, controlling for group fixed effects, as discussed in more detail below.

Altogether, we code 192 dyads as allies and 236 dyads as enemies. The remaining 5892 dyads are classified as neutral. Figure 2 illustrates the network of alliances and rivalries in the DRC. Not surprisingly, the FARDC have the highest centrality. In line with the narrative, the other groups with a high centrality are Rwanda, Uganda, and the main branches of the RCD. On average, a group has 2.95 enemies and 2.4 allies (see Table B.1).

<sup>10</sup>For the 15 actors with the greatest level of fighting involvement over the sample period, expert coding allows us to code 43 alliances (out of 101) and 13 enmities (out of 145).

In our baseline specification, we assume a time-invariant network, although we relax it in an extension. As discussed above, the system of alliances underwent major changes at the end of the First Congo War, while remaining relatively stable thereafter. The two main instances of changing relationships discussed in Section 3.1 involve the FARDC and Rwanda. Recall that ACLED splits the coding of Rwanda before and after 1999, and of the FARDC under Laurent and Joseph Kabila. This distinction is useful as it provides us with some flexibility in coding the two most important changes in the system of alliances. In particular, Rwanda-I is coded as an ally of Uganda and the RCD, while Rwanda-II is coded as an enemy of Uganda and of the RCD-K, and as an ally of the RCD-G. Similarly, we code the FARDC-LK and the FDLR as allies, and the FARDC-JK and the FDLR as neutral.<sup>11</sup> We test the robustness of the results to alternative assumptions (including merging Rwanda and the FARDC into two unified groups).

For the other dyads, we search for patterns in ACLED that may be suggestive of an inconsistent behavior, i.e., sometimes fighting together and sometimes against each other (see Appendix C for details). We detect eight such cases, and among them only two suggest the possibility of a switch in the nature of the link.<sup>12</sup> We deal with the problematic cases in the robustness analysis. While switching links appear to be rare, many groups are active only in few periods, and occasionally new groups are formed out of scissions of pre-existing groups. For this reason, in Section 4.2.1 we exploit an unbalanced sample where we allow for entry and exit of groups.

#### 4 ECONOMETRIC MODEL

Our empirical analysis is based on the model of Section 2.4 which allows for exogenous sources of heterogeneity in the OP of groups. Equation (11) can be estimated econometrically if one assumes that the individual shocks  $\tilde{\varphi}_i$  comprise both observable and unobservable components. More formally, we assume that  $\tilde{\varphi}_i = \mathbf{z}'_i \boldsymbol{\alpha} + \epsilon_i$ , where  $\mathbf{z}_i$  is a vector of group-specific observable characteristics, and  $\epsilon_i$  is an unobserved shifter. Replacing  $x_i$  and  $\varphi_i$  by their respective equilibrium values yields the following structural equation:

$$x_i^* = \varphi_i^*(G) - \beta \sum_{j=1}^n a_{ij}^+ x_j^* + \gamma \sum_{j=1}^n a_{ij}^- x_j^* - \mathbf{z}'_i \boldsymbol{\alpha} - \epsilon_i, \quad (13)$$

where we recall that  $\varphi_i^*$  is a function of the structural parameters  $\beta$  and  $\gamma$  and of the time-invariant network structure, while being independent of the realizations of individual shocks  $(\mathbf{z}_j, \epsilon_j)$  (see equation (4) and the analysis in Section 2.4). Our goal is to estimate the network parameters  $\beta$  and  $\gamma$ . The estimation is subject to a simultaneity or reflection problem, a common challenge in the estimation of network externalities. In this class of models, it is usually difficult to separate contextual effects, i.e., the influence of players' characteristics, from endogenous effects, i.e., the effect of outcome variables via network externalities. In our model, one might worry that omitted variables affecting  $x_i^*$  be spatially correlated, implying that one cannot safely assume spatial independence of  $\epsilon_i$ . Ignoring this problem would yield inconsistent estimates of the spillover parameters.

We tackle the problem through an IV strategy similar to Acemoglu *et al.* (2015). They study public good provision in a network of Colombian municipalities using historical characteristics

<sup>11</sup>This accords with our coding rule (no expert coding, multiple fights on the same and opposite camps). It is also consistent with the narrative that the FARDC has fought the FLDR more for tactical reason (i.e., to prevent Rwanda's direct intervention) than because of a deep hostility.

<sup>12</sup>In the other cases, there is no time pattern, and the volatile behavior appears to be the outcome of tactical fighting. This suggests a bilateral relation that is neither an alliance nor an enmity. Hence, coding the link as neutral seems accurate.

of local municipalities as instruments. In our case, it is difficult to find time-invariant group characteristics that affect the fighting efforts of a group’s allies or enemies without invalidating the exclusion restriction. For instance, cultural or ethnic characteristics of group  $i$  are likely to be shared by its allies. For this reason, we identify the model out of exogenous time-varying shifters affecting the fighting intensity of allies and enemies over time. This panel approach has the advantage that we can difference out any time-invariant heterogeneity, thereby eliminating the problem of correlated effects.

**Panel Specification** – We maintain the assumption of an exogenous time-invariant network, and assume the conflict to repeat itself over several years. We abstract from reputational effects, and regard each period as a one-shot game. These are strong assumptions which are necessary to make the empirical implementation feasible. The variation in conflict intensity over time is driven by the realization of group- and time-specific shocks, amplified or offset by the endogenous response of the groups, which, in turn, hinges on the network structure. More formally, we allow both  $x_i^*$  and  $\tilde{\varphi}_i$  to be time-varying.  $x_{it}^*$  corresponds to the annual number of ACLED events involving  $i$  in year  $t$ , and

$$\tilde{\varphi}_{it} = \mathbf{z}'_{it}\boldsymbol{\alpha} + e_i + \epsilon_{it}. \quad (14)$$

Here,  $\mathbf{z}_{it}$  is a vector of observable shocks with coefficients  $\boldsymbol{\alpha}$ ,  $e_i$  is an unobservable time-invariant group-specific shifter, and  $\epsilon_{it}$  is an i.i.d., zero-mean unobservable shock. Rainfall measures are examples of observable shifters  $\mathbf{z}_{it}$  that will be key for identification. The panel analogue of equation (13) can then be written as:

$$\text{FIGHT}_{it} = \text{FE}_i - \beta \times \text{TFA}_{it} + \gamma \times \text{TFE}_{it} - \mathbf{z}'_{it}\boldsymbol{\alpha} - \epsilon_{it}, \quad (15)$$

where  $\text{FIGHT}_{it} = x_{it}^*$  is group  $i$ ’s fighting effort at  $t$ ,  $\text{TFA}_{it} = \sum_{j=1}^n a_{ij}^+ x_{jt}^*$  is the total fighting effort of group  $i$ ’s allies,  $\text{TFE}_{it} = \sum_{j=1}^n a_{ij}^- x_{jt}^*$  is the total fighting effort of group  $i$ ’s enemies, and  $\text{FE}_i = +\varphi_i^*(G) - e_i$  is a group fixed effect capturing both the equilibrium OP level and unobservable time-invariant heterogeneity. This heterogeneity can be filtered out by including group fixed effects. However, the two covariates TFA and TFE are correlated with the error terms. Thus, OLS estimates are inconsistent due to an endogeneity bias.

**Instrumental variables (IV)** – The problem can be addressed by a panel IV strategy. Identification requires exogenous sources of variation in the fighting efforts of group  $i$ ’s allies and enemies that do not influence group  $i$ ’s fighting effort directly. To this aim, we use time-varying climatic shocks (rainfall) impacting the homelands of armed groups. In line with the empirical literature and historical case studies (Dell 2012), we focus on local rainfall as a time-varying shifter of OP, and hence the fighting effort of allies and enemies. More formally, our instruments are  $\text{RA}_{it} = \sum_{j=1}^n a_{ij}^+ \times \text{RAIN}_{jt}^*$  and  $\text{RE}_{it} = \sum_{j=1}^n a_{ij}^- \times \text{RAIN}_{jt}$ , where  $\text{RAIN}_{jt}$  denotes the rainfall in group  $j$ ’s territory. Take TFE, for instance. Above-average rainfall in the homelands of group  $i$ ’s enemies ( $\text{RE}_{it}$ ) reduces the enemy groups’ propensity to fighting because it increases agricultural productivity, thereby pushing up the reservation wages of local workers to be recruited by enemy armed groups. In other words, rainfall increases the opportunity cost of fighting. In addition, high rainfall could pose an obstacle to war activities (e.g., through mud roads), reinforcing the opportunity cost effect. These channels linking rainfall to conflict are in line with earlier studies (e.g., Miguel *et al.* 2004). There also potential offsetting effects: rainfall can increase revenues available to armed groups if agriculture is used as a source of taxation (see Fearon 2008). Our estimates below suggest that in our data this effect is dominated by the others.

To be a valid instrument, rainfall in the homelands of the allies (enemies) must be correlated with the allies' (enemies') fighting efforts. We document below that this is so in the data. In addition, rainfall must satisfy the exclusion restriction that rainfall in the homelands of group  $i$ 's allies and enemies has no direct effect on group  $i$ 's fighting effort. A first concern is that rainfall is spatially correlated, due to the proximity of the homelands of allied or enemy groups. However, this problem is addressed by controlling for the rainfall in group  $i$ 's homeland in the second-stage regression. For instance, suppose that group  $i$  has a single enemy, group  $k$ , and that the two groups live in adjacent homelands. Rainfall in  $k$ 's homeland is correlated with rainfall in  $i$ 's homeland. However, rainfall in  $k$ 's homeland is a valid instrument for  $k$ 's fighting effort, as long as rainfall in  $i$ 's homeland is included as a non-excluded instrument. A potential issue arises if rainfall is measured with error, and measurement error has a non-classical nature. We tackle this issue below in the robustness analysis.

Two additional threats to the exclusion restriction come from internal trade and migration. Rainfall may affect terms of trade. For instance, a drought destroying crops in Western Congo could cause an increase in the price of agricultural products throughout the entire DRC, thereby affecting fighting in the Eastern part of the country. Such a channel may be important in a well-integrated country with large domestic trade. However, inter-regional trade is limited in a very poor country like the DRC with a disintegrating government, very lacunary transport infrastructure, and a disastrous security situation. The war itself contributed to the collapse of internal trade, as documented by Zeender and Rothing (2010). The result is a very localized economy dominated by subsistence farming where spillovers through trade are likely to be negligible.

Weather shocks could trigger migration and refugee flows. For instance, an adverse weather shock hitting the homeland of one of group  $i$ 's enemies (say, group  $k$ ) could induce people to move from  $k$ 's to  $i$ 's homeland. The mass of displaced people could cause tensions and ultimately increase group  $i$ 's fighting for reasons other than changes in the fighting effort of  $k$  with a violation of the exclusion restriction. While we have no geolocalized statistical information to rule out this possibility, the aggregate evidence suggests that this is an unlikely issue. According to White (2014), the quasi-totality of migration movements in the DRC in the last decades have been caused by armed conflicts and concerns about security rather than by economic factors. For instance, only 0.7% of migrants indicate fleeing from natural catastrophes as the motivation for fleeing their homeland, while almost all refugees indicate that their movements are conflict- or security-related. It therefore appears very rare in the DRC that people are induced to migrate because of the scarcity of rain.<sup>13</sup> Similarly, rainfall could affect the activity of bandit groups. Yet, the boundaries between the activity of militias and bandit groups are thin in the DRC, so the activity of bandits is hardly a separate competing factor.

**Spatial Autocorrelation** – Since both violent events and climatic shocks are clustered in space, it is important to take into account spatial dependence in our data. For this reason, we estimate standard errors with a spatial HAC correction allowing for both cross-sectional spatial correlation and location-specific serial correlation, following Conley (1999 and 2008). However, there is no off-the-shelf application of these methods to panel IV regressions.<sup>14</sup> Therefore, we

---

<sup>13</sup>In addition, even people who are forced to move because of fighting “*reportedly try to stay close to home so that they can monitor their lands and track the local security situation.*” (White 2014: 6). This means that even people seeking shelter are unlikely to travel far away and penetrate zones of activity of other groups. An extended discussion of migration within the DRC can be found in Appendix C.

<sup>14</sup>See Vogelsang (2011) for an asymptotic theory for test statistics in linear panel models that are robust to heteroskedasticity, autocorrelation, and/or spatial correlation. Hsiang et al. (2011) provide a useful STATA code to calculate spatially correlated standard errors in panel regressions. Also, IV regressions are dealt with by Jeanty



program a STATA code that allows us to estimate Conley standard errors in a flexible fashion.<sup>15</sup> In the spatial dimension, we retain a radius of 150 km for the spatial kernel – corresponding to the 11th percentile of the observed distribution of bilateral distance between groups. More specifically, the weights in the covariance matrix are assumed to decay linearly with the distance from the central point of a group’s homeland, reaching zero after 150 km. We impose no constraint on the temporal decay for the Newey-West/Bartlett kernel that governs serial correlation across time periods. In other words, observations within the spatial radius can be correlated over time without any decay pattern. Robustness to alternative spatial and temporal kernels is explored in Appendix C.

A related challenge has been to adapt to our environment the test for weak instruments proposed by Kleinbergen and Paap (2006), henceforth, KP. KP is a rank test of the first-stage VCE matrix that is standardly used with 2SLS estimators and cluster robust standard errors. The statistic is valid under general assumptions, and the main requirement is that the first-stage estimates have a well-defined asymptotic VCE. To the best of our knowledge, the test had not been implemented in panel IV regressions with spatial HAC correction. We tackle a similar issue for the Hansen J overidentification test.

#### 4.1 ESTIMATES OF THE EXTERNALITIES

In this section, we estimate the regression equation (15) using a panel of 80 armed groups over 1998-2010. In all specifications, we include group fixed effects and year dummies, and estimate standard errors assuming spatial and within-group correlation as discussed above. In addition, all specifications control for current and lagged rainfall at the centroid of the group’s homeland, allowing for both a linear and a quadratic term.<sup>16</sup>

Table 1 displays the estimates of  $-\beta$  and  $+\gamma$  from second-stage regressions. Column 1 is an OLS specification. An increase in the enemies’ fighting effort (TFE) is associated with a higher fighting effort, consistent with the theory, whereas an increase in the allies’ fighting effort (TFA) has no significant effect. Since the OLS estimates are subject to an endogeneity bias, in the remaining columns we run a set of IV regressions. Column 2 replicates the specification of column 1 in a 2SLS setup using the lagged fighting efforts of each group’s set of enemies and allies as excluded instruments. In accordance with the predictions of the theory, the estimated coefficients of TFE and TFA are positive (0.13) and negative (-0.22), respectively, and statistically significant at the 5% level.

The associated first-stage regressions are reported in the corresponding columns of Table 2, where, for presentational purposes, only the coefficients of the excluded instruments are displayed. It is reassuring that the lagged rainfall in the enemies’ homelands has a negative effect on the enemies’ (while not on the allies’) fighting effort, whereas the lagged rainfall in the allies’ homelands has a negative effect on the allies’ (while not on the enemies’) fighting effort. This pattern, which conforms with the theoretical predictions, is confirmed in all specifications of Table 2. The KP-stat of 10.6 raises a (mild) concern about weak instruments, an issue to which we return below.

In the parsimonious specification of column 2, the coefficients of interest may be affected by some time-varying shocks that affect asymmetrically the armed groups’ incentives to fight. For instance,

---

(2012). However, neither routine handles spatial correlation in panel IV regressions.

<sup>15</sup>We owe a special thanks to Rafael Lalive for his generous help in this task.

<sup>16</sup>As discussed in Section 3.2, following the ACLED coding, the FARDC and Rwanda are each split in two groups according to the period of activity. When a group is inactive (e.g., the FARDC-JK during 1997-2001), its fighting effort is set equal to zero. To avoid that these artificial zero observations affect the estimates of the structural parameters  $\beta$  and  $\gamma$ , we always include in the regressions a group dummy interacted with a full set of year dummies for the period of inactivity. In Section 4.2, we show that the estimates are robust to merging Rwanda and the FARDC into a unique group each.

Table 1: Baseline regressions (second stage).

Dependent variable: Total fighting								
	OLS	Reduced IV	Full IV	Neutrals	Battles	$(d^-, d^+)$ $\geq 1$	GED coord.	GED union
Total Fight.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Enemies (TFE)	0.066*** (0.016)	0.130** (0.057)	0.066*** (0.019)	0.083*** (0.019)	0.081*** (0.020)	0.091*** (0.022)	0.084*** (0.019)	0.125*** (0.039)
Allies (TFA)	0.001 (0.017)	-0.218** (0.086)	-0.117*** (0.035)	-0.114*** (0.033)	-0.117*** (0.037)	-0.157*** (0.058)	-0.112*** (0.032)	-0.117*** (0.036)
Neutrals (TFN)				0.004 (0.004)	0.004 (0.005)	0.013 (0.013)	0.004 (0.004)	0.006 (0.004)
Add. controls	Reduced	Reduced	Full	Full	Full	Full	Full	Full
Estimator	OLS	IV	IV	IV	IV	IV	IV	IV
Instrum. Var.	n.a.	Restricted	Full	Full	Full	Full	Full	Full
Kleibergen-Paap	n.a.	10.6	19.5	22.5	20.6	17.8	22.1	10.4
Hansen J	n.a.	0.16	0.68	0.58	0.53	0.66	0.58	0.69
Observations	1040	1040	1040	1040	988	598	1040	1781
R-squared	0.510	0.265	0.579	0.568	0.567	0.537	0.569	0.516

Note: The unit of observation is an armed group in a given year. The panel contains 80 armed groups between 1998 and 2010. All regressions include group fixed effects and control for rainfall in the group's homeland. Columns 1–3 include time fixed effects. Robust standard errors corrected for spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 2: Baseline regressions (first stage).

Dep. Variable:	IV regress. of col. (2)		IV regress. of col. (3)		IV regress. of col. (4)	
	TFE (1)	TFA (2)	TFE (3)	TFA (4)	TFE (5)	TFA (6)
<b>Rain (<math>t - 1</math>) Enem.</b>	<b>-1.595***</b> <b>(0.297)</b>	-0.019 (0.141)	<b>-1.354***</b> <b>(0.332)</b>	0.277* (0.156)	<b>-1.327***</b> <b>(0.322)</b>	0.291** (0.139)
Sq. Rain ( $t - 1$ ) Enem.	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)
<b>Rain (<math>t - 1</math>) All.</b>	0.126 (0.283)	<b>-0.929***</b> <b>(0.155)</b>	0.028 (0.222)	<b>-0.588***</b> <b>(0.192)</b>	0.089 (0.219)	<b>-0.571**</b> <b>(0.225)</b>
Sq. Rain ( $t - 1$ ) All.	-0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)
<b>Current Rain Enem.</b>			<b>-1.125***</b> <b>(0.243)</b>	0.131 (0.102)	<b>-0.936***</b> <b>(0.257)</b>	0.073 (0.108)
Sq. Curr. Rain Enem.			0.000*** (0.000)	-0.000*** (0.000)	0.000* (0.000)	-0.000** (0.000)
<b>Current Rain All.</b>			-0.461** (0.204)	<b>-0.366***</b> <b>(0.123)</b>	-0.414** (0.210)	<b>-0.448***</b> <b>(0.164)</b>
Sq. Curr. Rain All.			0.000 (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)
Kleibergen-Paap F-stat	10.6	10.6	19.5	19.5	22.5	22.5
Hansen J (p-value)	0.16	0.16	0.68	0.68	0.58	0.58
Observations	1040	1040	1040	1040	1040	1040

Note: The unit of observation is an armed group in a given year. The panel contains 80 armed groups between 1998 and 2010. All regressions include group fixed effects and control for rainfall in the group's homeland. Columns 1-4 contain time fixed effects. Robust standard errors corrected for spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

global economic or political shocks may change the pressure from international organizations, which in turn affects mainly the war activity of foreign armies, government organizations, or more generally of large combatant groups. To filter out such time-varying heterogeneity, in columns 3–8 we control for three time-invariant characteristics (*Government Organization*, *Foreign*, and *Large*) interacted with a full set of year dummies.<sup>17</sup> The description of these three variables can be found in Section 3.2 above. Together with adding control variables, we expand the set of excluded instruments (i.e., the rainfall measures), in order to improve the predictive power of the first-stage regression.<sup>18</sup> The expanded set of instruments now comprises current-year and lagged-year rainfall (with a linear and a quadratic term) of allies and enemies, as well as current and lagged rainfall of degree-two neighbors (i.e., enemies’ enemies and allies’ enemies), both with a linear and a quadratic term.<sup>19</sup>

The estimated coefficients in column 3 continue to feature the alternate sign pattern predicted by the theory. Their magnitude is smaller than in column 2, but the coefficients are estimated more precisely, being statistically significant at the 1% level. In column 4, we add to the vector of regressors TFN (Total Fighting of Neutrals), which is defined in analogy with TFA and TFE. Hence, we now also add to the set of instruments the current and lagged rainfall in the territories of neutral groups (both as a linear and quadratic term). Since the theory predicts that the exogenous variation in TFN should have no effect on the dependent variable, this is a useful test of the theoretical predictions. The prediction is borne out in the data: the point estimate of TFN is very close to zero and statistically insignificant. The first-stage regressions yield large KP-stats (19.5 in column 3 and 22.5 in column 4), suggesting no weak instrument problem. Column 4 is our preferred specification and will be the basis of our robustness checks in the following sections below.

Our measure of fighting intensity is coarse insofar as it does not weigh events by the amount of military force involved. Ideally, we would like to have information about the number of casualties or other measures of physical destruction. However, this information is available only for very few events. This raises the concern that the results may be driven by small events (e.g., local riots or minor skirmishes). As discussed in Section 3.2, ACLED distinguishes between different categories of events. In column 5, we measure fighting effort in a more restrictive fashion, by only counting events that are classified in ACLED as battles. This addresses two issues: first, battles are less likely to get unreported by media; second, it would be reassuring to see that the estimates of  $\beta$  and  $\gamma$  are robust to excluding small events that represent a share of 42% of total events. The estimated coefficients are indeed very similar when we use only information on battles, with no evidence of weak instruments (KP=20.6).<sup>20</sup>

A related concern is that many of the 80 groups are involved only in a small number of events. Although heterogeneity in group size is controlled for by fixed effects, one might be concerned that the estimation of the externalities hinges on the occasional operation of small groups. In lack of a direct measure of group size, in column 6 we restrict the analysis to the 46 groups that have at least one friend and one enemy, proxying for being relatively important actors. This restriction reduces

---

<sup>17</sup>In Section 4.2.3 below, we find that some results are sensitive to the threshold used to construct the dummy variable *Large*. For this reason, in Table B.2 we replicate Table 1 when *Large* is replaced by *Large\_6* (see Section 3.2). The results are robust.

<sup>18</sup>We also run the specification of column 2 with the expanded set of instruments. The estimated coefficients of interest are 0.15 for TFE (s.e. 0.05) and -0.15 for TFA (s.e. 0.06). The KP-test yields the value 19.7.

<sup>19</sup>When we use the current and past average rainfall in enemies’ and allies’ homelands as instruments, we also control for the current and past average rainfall in the group’s homeland in the second-stage regression. This is important, since the rainfall in enemies’ and allies’ homelands is correlated with the rainfall in the group’s homeland. Omitting the latter would lead to a violation of the exclusion restriction. The results are robust to including further instruments, for instance, the allies’ allies and the enemies’ allies.

<sup>20</sup>The number of observation falls to 988, as 4 of the 80 groups drop out of the sample for never being involved in any battles.

the network size, causing a 40% drop in the number of observations. Reassuringly, the estimated externalities are larger than in column 4 ( $\beta = 0.16$  and  $\gamma = 0.09$ ). The KP-stat is 17.8.

The accuracy of the geolocalization in ACLED has been questioned, as discussed in Section 3.2 above. For this reason, we integrate ACLED with information from the GED, which has been argued to be more accurate in this respect. We cannot simply replace ACLED with GED data because (i) the number of observations would drop by two thirds, aggravating underreporting concerns; (ii) for each event, GED lists at most one group on each side of the clash. However, in 1090 cases it is possible to match events in GED and ACLED beyond reasonable doubt. In these cases, we use the geolocalization in GED to identify the groups' homelands. For the events that cannot be matched, we continue to use the geolocalization in ACLED. The results, provided in column 7, are indistinguishable from those in column 4 (with KP=22.1).

In addition, we use the union set of the events in GED and ACLED, i.e., we construct a larger dataset that merges the matched events with all unmatched events in either dataset. By this procedure, the number of fighting events increases from 4676 to 5078.<sup>21</sup> There is also a larger number of armed groups, 137 instead of 80. This procedure involves some heroic assumptions, and is subject to the risk that our algorithm fails to match some events that are in fact reported by both datasets, thereby causing an artificial duplication of events. With this *caveat* in mind, we find the estimates of TFE, TFA, and TFN to be, respectively, positive, negative, and insignificant, in accordance with the theory. The order of magnitude of the coefficients is comparable with those in column 4, and the point estimates are in fact larger in absolute value. However, the KP-stat is now lower (10.4). The details of the constructions of the merged dataset are in Appendix C.

The externalities are quantitatively large. Consider the estimates in column 4. The average number of yearly events in which a group is involved is 6, and its standard deviation is 25. Hence, a one standard deviation increase in TFE (i.e., 110 events) translates into a 0.37 increase in total fighting (i.e., 9 events). A one standard deviation increase in TFA (i.e., 86 events) translates into a 0.39 decrease in total fighting (i.e., 10 events). An estimate of the global effect of the network externalities is provided in Section 5.3 below.

We have also checked that, conditional on the estimates of  $\beta$  and  $\gamma$ , condition (3) holds true for *all* groups in conflict in all IV specifications of Table 2. Finally, the null hypothesis of the Hansen J test is not rejected in any specification, indicating that the overidentification restrictions are valid.

## 4.2 ROBUSTNESS ANALYSIS

We run a large battery of robustness checks. here, we summarize the most important ones.

### 4.2.1 Variation over Time in the Network Structure

In our dataset, many groups are not active in all periods. We also observe new groups entering the conflict at a later stage, and a few groups which stop fighting. While in the analysis of Section 4.1 we interpret zero fighting events as a low fighting effort, the absence of armed engagements could alternatively indicate that a group does not take part in the conflict in a particular subperiod. For this reason, in the first robustness check we address this concern by recognizing that the number of groups that are in the network can change over time.

---

<sup>21</sup>In particular, of the 1641 groups in GED, 402 are very likely missing in ACLED, 1090 can be accurately matched with very high probability to ACLED events, while 149 events are likely to refer to given events already in ACLED but for which the match cannot be proven with high enough probability. Hence, in column 8 we follow the conservative approach of only adding the GED events missing with very high probability in ACLED (i.e., 402 additional events) to avoid double counting.

Table 3: Time-varying network.

	Dependent variable: Total fighting						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Total Fight. Enemies (TFE)	0.085*** (0.022)	0.074*** (0.023)	0.088*** (0.024)	0.138*** (0.031)	0.075 (0.048)	0.068** (0.030)	0.211*** (0.047)
Total Fight. Allies (TFA)	-0.115*** (0.031)	-0.097*** (0.025)	-0.106*** (0.030)	-0.212*** (0.065)	-0.143** (0.065)	-0.128*** (0.041)	-0.251*** (0.070)
Total Fight. Neutrals (TFN)	0.006 (0.005)	0.003 (0.004)	0.002 (0.005)	0.048** (0.023)	0.022 (0.021)	0.006 (0.010)	-0.022** (0.009)
Kleibergen-Paap F-stat	27.1	8.7	11.7	9.2	5.5	15.3	n.a.
Hansen J (p-value)	0.61	0.48	0.60	0.71	0.74	0.51	n.a.
Observations	1040	1040	1040	469	322	637	1040
R-squared	0.603	0.634	0.594	0.501	0.627	0.594	0.179

Note: An observation is a given armed group in a given year. The panel contains 80 armed groups between 1998 and 2010. All regressions include group fixed effects and the full set of controls and instruments (like in baseline column 4 of Table 2). Columns 1-3 define windows of activity and include a group-specific dummy for periods when a group is inactive. In column 1, inactivity is defined by expert coding combined with ACLED information. In column 2, inactivity is defined based on ACLED information only. In column 3, inactivity is based on ACLED information + or - 3 years. Columns 4-6 implement an ILLE estimator on the unbalanced sample of active groups only using the same windows of activity as in columns 1-3. Column 7 performs an instrumented Tobit based on a control function approach. Cluster robust standard errors are corrected for spatial HAC in columns 1-6 and are bootstrapped in column 7. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

We use a variety of expert sources to check when each group started its activity, and when, if at all, it ceased to be militarily active. We could gather information for 38 groups (many of them being active in the entire period). However, no official date of establishment or disbandment is available for informal organizations such as ethnic militias. For these groups, we construct a window  $[S - \omega, T + \omega]$ , where  $S$  and  $T$  are, respectively, the first and last year in which we see the group being active (i.e.,  $x_{it} > 0$ ). We add a window of  $\omega \geq 0$  since the groups might have existed prior to their first or after their last recorded engagement. The details of the construction of the dataset are provided in Appendix C.

We estimate the model by the following three strategies:

1. We add to the baseline specification a set of group-specific dummies switching on in all periods in which the group is suspected to be inactive.
2. We adjust, in addition, the estimation procedure to make it fully consistent with the structural model. To see why, consider equations (4)–(6). When the number of groups in the network changes over time, one must replace  $\varphi_i^*(G)$  by a time-varying analogue given by

$$\varphi_{i,t}^*(G; \beta, \gamma) = \left(1 - \frac{1}{\sum_{i=1}^{n,t} \Gamma_{i,t}^{\beta,\gamma}(G)}\right) \left(\frac{1}{\sum_{i=1}^n \Gamma_{i,t}^{\beta,\gamma}(G)}\right) \Gamma_{i,t}^{\beta,\gamma}(G),$$

where  $\Gamma_{i,t}^{\beta,\gamma}(G) = 1 / (1 + \beta d_{i,t}^+ - \gamma d_{i,t}^-)$ . When  $\varphi_{i,t}^*$  is time-varying, it is no longer absorbed by the group fixed effects. However, the model can still be estimated. In particular, one can then estimate the following regression equation:

$$\text{FIGHT}_{it} = \text{FE}_i + \varphi_{i,t}^*(G; \beta, \gamma) - \beta \times \text{TFA}_{it} + \gamma \times \text{TFE}_{it} - \mathbf{z}'_{it} \boldsymbol{\alpha} - \epsilon_{it}. \quad (16)$$

Here,  $\varphi_{i,t}^*$  can be estimated conditional on a prior for  $\beta$  and  $\gamma$ , as  $d_{i,t}^+$  and  $d_{i,t}^-$  are observable for

all  $i$  and  $t$ . Thus, we implement the Iterated Linear Least-square Estimator (ILLE) developed by Blundell and Robin (1999).<sup>22</sup>

3. We estimate the model using instrumented Tobit based on a control function approach.

Table 3 displays the results. All columns report analogues of column 4 in Table 1. Columns 1–3 correspond to the first approach. In column 1, the time window is set to  $\omega = 0$ ; in column 3, we set  $\omega = 3$ ; in column 2, we code as a period of possible inactivity any consecutive spell of zeros at the beginning or at the end of the sample, using only the information from ACLED. The estimates of  $\beta$  and  $\gamma$  are similar to those in column 4 in Table 1. The KP-stats are 25.3, 8.7, and 11.7, respectively. Columns 4–6 correspond to the second approach. In spite of a drastic sample size reduction, the coefficients continue to have the same order of magnitude as in the baseline table. In column 4, the coefficients are larger in absolute value, and the coefficient of TFN turns significant, while remaining much smaller than those of TFE and TFA. In column 5, the coefficient of TFE turns insignificant. In column 6, the results are very similar to column 4 in Table 1. The KP-stats are 6.5, 5.5, and 15.3, respectively. The weak instruments in columns 4 and 5 are not surprising, since the number of observations is, respectively, one third and one half of that in the full sample. Also, this specification is very demanding, since in many cases no reported involvement in ACLED events may indicate a low level of fighting activity rather than an outright withdrawal from the conflict. Column 7 is based on Tobit with a control function approach for the two-stage instrumentation. The estimated coefficients have the usual alternate sign pattern, but are now much larger in absolute value. Overall, we find these results reassuring.

#### 4.2.2 Alternative Specifications

In this section, we consider three sets of robustness checks. All regression tables can be found in Appendix B.

**Second-Degree IV, Salient Events, and Alternative Network Construction:** A miscellany of important robustness checks is summarized in Table B.3. In column 1, we use only the rainfall in the homeland of degree-two neighbors (e.g., the rain of enemies’ enemies and of allies’ enemies) as excluded instruments, following Bramoullé *et al.* (2009).<sup>23,24</sup> In column 2, we use the information for the subperiod 1998-2002 to estimate the network links, and the panel for 2003-10 to estimate the spillover coefficients. In columns 3-4, we restrict attention to salient episodes for which measurement error is likely to be less important. In column 3, we drop all events with zero fatalities (while keeping events for which the number is unknown). In column 4, we restrict

---

<sup>22</sup>We start by guessing  $(\beta_0, \gamma_0)$  and  $\varphi_{i,t}^*(G; \beta_0, \gamma_0)$ . Then, we obtain a first set of estimates  $(\hat{\beta}_1, \hat{\gamma}_1)$  conditional on the guess, update  $\varphi_{i,t}^*(\hat{\beta}_1, \hat{\gamma}_1)$ , and re-estimate the model iteratively until we converge to a fixed point. Computationally, we stop the iteration as soon as  $\|(\hat{\beta}_n, \hat{\gamma}_n) - (\hat{\beta}_{n-1}, \hat{\gamma}_{n-1})\| < 0.0001$  (i.e., two orders of magnitude smaller than the estimated standard errors). While Blundell and Robin (1999) address the issue of endogenous regressors with a control function approach (i.e., first stage estimated residuals included as regressors in the second stage), we iterate on our 2SLS estimator that accommodates spatially clustered robust standard errors. We checked that the control function-ILLE and 2SLS-ILLE yield identical point estimates.

<sup>23</sup>In particular, we continue to treat as excluded instruments the rainfall in the enemies’ enemies’ homelands, the rainfall in the allies’ enemies’ homelands, and the rainfall in the neutrals’ homelands. However, the rainfall in the enemies’ homelands and the rainfall in the allies’ homelands are treated as control variables. For all rainfall measures, we take the linear and square term and the current rain, first lag, and second lag.

<sup>24</sup>Note that, contrary to their model, in our theory there is no reason why an instrumentation based on first-order links should yield inconsistent estimates. As discussed above, the case for our regressions to be contaminated by contextual effects is weak in our panel regression.

attention to battles, riots, and violent events. In column 5, we exclude all events involving group  $i$  when computing the total fighting efforts of allies and enemies of group  $i$ . For example, if the LRA’s enemies are involved in 10 clashes in the year 2000, and 3 of them involve the LRA, then the measure of TFE used in the regression would take the value of 7. In column 6, we control for the lagged total fighting effort of both enemies and allies. In columns 7-8, we test the robustness of the results to different definitions of enmities and alliances: in column 7, we code two groups as enemies if they have been observed clashing on at least *one* occasion, and if they have never been observed co-fighting on the same side; in column 8, we only code two groups as allied if they have been observed co-fighting on at least *two* occasions during the sample period and if they have had less than two clashes, while analogously we code them as enemies if they have fought on at least two occasions against each other, while having had less than two co-fighting incidents. Further, in column 9, we instrument the network links with dyadic characteristics (co-ethnicity, spatial proximity of group centroids, etc.). The observed links are replaced by probabilities of link formation as predicted by a random utility model discussed in Section 6.2 below. Finally, in column 10 we include in the second stage regression the cubic terms of past and current rain.

In terms of results, the coefficient of TFE is always positive, highly significant, and stable. Likewise, the coefficient of TFA is always negative and significant with the exception of column 2. In most cases, the KP-stat is above 20. The coefficient of TFN is always very close to zero and insignificant.

**Group Definition (FARDC, Rwanda, & Others):** In our benchmark analysis, we have followed the rule of treating groups as separate entities whenever they are classified as such by ACLED. This agnostic way of proceeding has the advantage of not requiring any discretionary coding decision. We check the robustness of our results in this dimension.<sup>25</sup> In Table B.4, we show that the results are robust to (i) treating the FARDC-LK and FARDC-JK as one single actor; (ii) merging all local Mayi-Mayi militia branches into one single actor; (iii) merging Rwanda-I and Rwanda-II into a single group; (iv) treating both the FARDC and Rwanda as two single actors.

**Ambiguous Network Links:** Table B.5 deals with ambiguous network links, i.e., links where the narrative might suggest different coding than the one we used. First, we consider the fragile relationship between Uganda and Rwanda (see Section 3.1 for historical background). In our baseline regression, our coding rule classifies Rwanda-I and Uganda as allies (until 1999), whereas Rwanda-II and Uganda are coded as enemies (after 1999). In columns 1 and 2, we code Rwanda and Uganda as always neutral and always allies, respectively. In column 3, instead, we code them as allies until 1999, and as neutral thereafter. Next, we consider another ambiguous relationship, i.e., the FARDC *vs.* the FDLR. In the baseline estimates, they are first allies (until 2001), and then neutral. Here, we assume that they are enemies after 2001 (column 4), or neutral throughout the entire period (column 5). Other ambiguous network links are discussed in Appendix B. The results are robust.

#### 4.2.3 Alternative Rainfall Data

A concern with our IV strategy is that the rainfall variable may be subject to measurement error. The GPCC data are based on interpolation on information coming from a limited number of gauges in the DRC and neighboring countries. The ensuing classical measurement error could attenuate the power of the rainfall instruments. A potentially more severe problem might arise if the gauge-based

---

<sup>25</sup>In all the robustness checks of Table B.4, we re-estimate the network for each of the different specifications.

measure systematically underreports precipitations in war zones, e.g., because fighting destroys gauges.

To address this concern, we turn to satellite-based rainfall estimates (tables are in the appendix). These are not affected by the dynamics of conflict although they are typically noisier than gauge-based measures, especially when one look at the within-cell variation over time.<sup>26</sup> We consider two satellite-based data series: GPCP and TRMM. GPCP is a well-established data source which has been used in several existing studies (e.g., Miguel *et al.* 2004, and Brückner and Ciccone 2011.) However, it is only available at a  $2.5 \times 2.5$  resolution, which is considerably more aggregate than GPCP. TRMM has a finer ( $0.25 \times 0.25$ ) resolution but the data quality is more controversial (see Appendix C), especially in the within-cell time variation, which is our source of identification.

Given these premises, we make a twofold use of the satellite data. First, we use them to assess whether gauge-based measures are biased and suffer from non-classical measurement error. Second, we run regressions directly on GPCP and TRMM data. For the first purpose, consider the following simple model:

$$\begin{aligned} \text{RAIN}_{ct}^{\text{sat}} &= \psi_c^{\text{sat}} + \text{RAIN}_{ct} + v_{ct}^{\text{sat}}, \\ \text{RAIN}_{ct}^{\text{gau}} &= \psi_c^{\text{gau}} + \text{RAIN}_{ct} + \tilde{v}_{ct}^{\text{gau}}, \end{aligned} \quad (17)$$

where  $c$  denotes the grid-cell at which rainfall is measured,  $\text{RAIN}_{ct}$  is the true (unobservable) rainfall, and  $v_{ct}^{\text{sat}}$  and  $\tilde{v}_{ct}^{\text{gau}}$  are the measurement errors.  $v_{ct}^{\text{sat}}$  is assumed to be i.i.d. The error term of the gauge measure is potentially subject to violence-driven measurement error. This possibility is allowed by letting  $\tilde{v}_{ct}^{\text{gau}} = \xi \times \text{VIOLENCE}_{ct} + v_{ct}^{\text{gau}}$ , where  $v_{ct}^{\text{gau}}$  is an i.i.d. error term. One can eliminate  $\text{RAIN}_{ct}$  from the previous system of equations and obtain:

$$\text{RAIN}_{ct}^{\text{gau}} = \psi_c + \text{RAIN}_{ct}^{\text{sat}} + \xi \times \text{VIOLENCE}_{ct} + \nu_{ct}, \quad (18)$$

where  $\psi_c = \psi_c^{\text{gau}} - \psi_c^{\text{sat}}$  and  $\nu_{ct} = v_{ct}^{\text{gau}} - v_{ct}^{\text{sat}}$  are, respectively, a grid-cell fixed effect and an i.i.d. disturbance. Our null hypothesis is that  $\xi = 0$ . If  $\xi \neq 0$ , the gauge-based measure suffers with non-classical measurement error.

We run a regression based on equation (18), measuring violence by the number of conflicts in ACLED. Table B.7 summarizes the results. Columns 1–4 report the results when satellite-based rainfall measures are retrieved from TRMM. Column 1 is a cross-sectional specification; column 2 includes grid-cell fixed effects – consistent with equation (18). In columns 3 and 4, we consider a log-linear specification where the two rainfall measures are log-scaled; this corresponds to a multiplicatively separable specification of model (17). Finally, we replicate the same set of four specifications in columns 5–8 with the GPCP satellite measure. Year dummies are included in all regressions. Standard errors are clustered at the grid-cell level.<sup>27</sup>

As expected, there is a highly significant correlation between the gauge- and the satellite-based rainfall measures. Interestingly, the within-cell correlation between GPCP and satellite data (65% with GPCP and 61% with TRMM) is significantly smaller than the overall correlation. The correlation between GPCP and TRMM (52%) is even lower. In no case are the estimates of  $\xi$  different from zero, with point estimates switching signs across specifications. The hypothesis

<sup>26</sup>In a recent survey on the new climate-economy literature, Dell *et al.* (2014, p.750) write: "Satellite data are not necessarily a panacea... [they] do not directly measure temperatures or precipitation, but rather make inferences from electromagnetic reflectivity in various wave length bands. A satellite-based series is not drawn from a single satellite, but rather from a series of satellite. Sensors have changed subtly over the years and, within a particular satellite, corrections are needed due to subtle changes in the satellite's orbit over time and other factors."

<sup>27</sup>Recall that the GPCP satellite measure is only available at the  $2.5 \times 2.5$  degree level, i.e., for larger cells than the two other measures that are at the  $0.5 \times 0.5$  degree level. In this case, we cluster at the  $2.5 \times 2.5$  cell level.



that  $\xi$  is negative due to the destruction of gauges in battlefields is strongly rejected, especially in specifications with grid-cell fixed effects. We conclude that there is no evidence that the gauge-based GPCC precipitation data are subject to non-classical measurement error in the DRC.

Next, we run the main specification in Table 1 using satellite rainfall data instead of GPCC data as instruments. Table B.8 shows the results. Since the point estimates are sensitive to the inclusion and the definition of the control variable *Large*, we present six columns for each data source. Each pair of columns corresponds to Columns 4–5 in Table 1, with the following differences: Columns 1–2 uses *Large*, Columns 3–4 uses *Large\_6*, and Columns 5–6 exclude *Large* altogether from the set of control variables. We aggregate the TRMM data at the 0.5 level and replace the missing observation for 1998 with the corresponding observation from the GPCP. Panel *a* shows the results. The coefficients in this panel follow the pattern of columns 4–5 in Table 1. They have the expected alternate sign, and are in most cases statistically significant, although some are smaller in absolute value. Panel *b* shows the results using GPCP. All coefficients have the expected signs but the coefficients of TFE are in some cases small and insignificant. The low KP-Stats suggest a potential weak instrument problem. This is not unexpected given the higher level of aggregation of the GPCP. Finally, panel *c* shows the result of regressions where we include *both* TRMM and GPCP rainfall data as instruments. This is a reasonable choice, since the time-series correlation between the two series is rather low, implying that there is signal in including both measures. The effects are consistent with those in panels *a–b*, although they are more precisely estimated.

In summary, the regressions using satellite data as instruments broadly confirm those from GPCC but are more fragile. We suspect that this is due to the coarser aggregation in the GPCP data, and to the noisier nature of the within-cell variation of TRMM, in line with the meteorology literature discussed in Appendix C.

#### 4.2.4 Measurement Error

Another concern here is that the network may be measured with error. Chandrasekhar and Lewis (2016) show that measurement error of network links can give rise to inconsistent estimates. Moreover, the bias can work in different directions, and there is no general recipe to correct it. To address this issue, we follow a Monte Carlo approach based on rewiring links in the observed network at random, and measuring the robustness of our estimates in such perturbed networks. We consider different assumptions about the extent and nature of measurement error of the network. All tables are in the appendix.

In our first exercise, we postulate a data generating process, and then we introduce a model of mismeasurement of network links. Then, we estimate the model as if the econometrician did not know the true network, but had to infer it from data measured with error. This procedure is generated for 1,000 realizations of measurement error. The results are reported in Table B.9 and the procedure is explained in more detail in Appendix C. The general lesson from this exercise is twofold. First, the Monte Carlo-generated measurement error in the links leads to an attenuation bias. Thus, if some information about existing links is missing, our regression analysis underestimates the spillover effects. Second, the extent of the bias is quantitatively modest. A measurement error of the order of 10% (which we regard as fairly large) yields an underestimate of the spillover parameters of 12% for  $\beta$  and 23% for  $\gamma$ .

In our second exercise, instead of assuming that some network links are missed, we bootstrap the entire analysis. In each Monte Carlo trial, first, we sample the ACLED events with replacement; then, we construct fake networks based on expert coding supplemented with the ACLED resampled events; then, we rebuild our main outcome variable,  $\text{FIGHT}_{it}$ , the explanatory variables  $\text{TFA}_{it}$  and

TFE<sub>it</sub>, and the related rainfall measures; finally we estimate all the specifications of the baseline Table 1 and store the relevant point estimates. This procedure is repeated 1,000 times in order to get the sampling distribution of each coefficient of interest. Table B.10 displays the results. The rows report the mean/standard errors of the sampling distribution for each coefficient of interest across the columns of Table 1. Visual inspection confirms that the bootstrap results are very similar to the baseline ones.

Finally, we check the effect of measurement error that is correlated with the group size. In particular, ACLED could systematically underreport events involving small groups. Although events involving small groups are also likely to be less significant, it is useful to study this possibility. We proceed by removing from the ACLED dataset events involving small groups (i.e., with a total number of enemies and allies strictly below the median) with a given binomial probability that is constant across Monte Carlo trials. Then, we operate as above by considering the fake ACLED dataset for coding the network, the IVs, etc. Table B.11 shows that there is a very mild attenuation bias. Even when 50% of times the participation of small groups goes unrecorded, the estimated coefficients of TFA, TFE, and TFN are very close to our baseline (col. 4 of Table 1) estimates. We conclude that under measurement of events involving small groups is not a major problem.

## 5 POLICY INTERVENTIONS

In this section, we perform counterfactual policy experiments. First, we consider interventions that selectively induce some fighting groups to exit the contest. Next, we consider policies (such as an arms embargo) that increase the marginal cost of fighting for selected groups. Finally, we study the effect of pacification policies, where enmity links are selectively turned into neutral ones.

The analysis is based on the simulation of counterfactual equilibria. To this aim, let  $G^b$  denote the benchmark network in which all groups fight. We maintain  $V = 1$ , and set the externality parameters equal to their baseline point estimates  $\hat{\beta} \approx 0.083$  and  $\hat{\gamma} \approx 0.114$  (column 4, Table 1).<sup>28</sup> Since the parameters are estimated with error, and the point estimates vary to some extent across specifications, the results must be taken with a grain of salt. We view the analysis as a useful illustration of the quantitative predictions of the theory as well as a tentative indication of the effects of alternative policy interventions in the Congo conflict.

The equations (6) and (13)–(15) allow us to estimate  $e_i$ , the time-invariant unobserved heterogeneity. More formally:

$$\hat{e}_i = -\widehat{\text{FE}}_i + \Lambda^{\hat{\beta}, \hat{\gamma}}(G^b) \left(1 - \Lambda^{\hat{\beta}, \hat{\gamma}}(G^b)\right) \Gamma_i^{\hat{\beta}, \hat{\gamma}}(G^b), \quad (19)$$

where  $\widehat{\text{FE}}_i$  is the estimated group fixed effect,  $\Gamma_i^{\hat{\beta}, \hat{\gamma}}(G^b) = 1/(1 + \hat{\beta}d_i^+ - \hat{\gamma}d_i^-)$ , and  $\Lambda^{\hat{\beta}, \hat{\gamma}}(G^b) = 1 - 1/(\sum_j \Gamma_j^{\hat{\beta}, \hat{\gamma}}(G^b))$ . We collapse the vector of time-varying shifters  $\mathbf{z}_{it}$  (rainfall, etc.) to its sample average,  $\bar{\mathbf{z}}_i = \sum_{t=1998}^{2010} \frac{\mathbf{z}_{it}}{13}$ , and denote by  $\bar{\mathbf{Z}} = \{\bar{\mathbf{z}}_i\}$  the estimated matrix of shifters. In other words, we compare an average year of conflict in the benchmark model to its corresponding counterfactual. We consistently set the time-varying i.i.d. shocks  $\epsilon_{it}$  to zero for all groups.

Following the analysis in Section 2.4, the vector of (Nash) equilibrium fighting efforts is obtained by inverting the system of equilibrium conditions implied by equations (13) and (14). In matrix form, this yields:

$$\mathbf{x}^*(G^b) = (\mathbf{I} + \hat{\beta}\mathbf{A}^+(G^b) - \hat{\gamma}\mathbf{A}^-(G^b))^{-1} \left[ \Lambda^{\hat{\beta}, \hat{\gamma}}(G^b)(1 - \Lambda^{\hat{\beta}, \hat{\gamma}}(G^b))\mathbf{\Gamma}^{\hat{\beta}, \hat{\gamma}}(G^b) - (\bar{\mathbf{Z}}\hat{\boldsymbol{\alpha}} + \hat{\mathbf{e}}) \right]. \quad (20)$$

<sup>28</sup>All second-order conditions (cf. equation (3)) continue to hold for all groups in the counterfactual experiments in which one player is removed.

Based on this equilibrium, we evaluate the effects of unanticipated policy shocks that affect either the network  $G^b$  or some exogenous parameters. We measure the welfare effects by the counterfactual changes in rent dissipation as defined in equation (10).

## 5.1 REMOVING ARMED GROUPS

Consider a policy intervention that induces some groups to leave the contest. Formally, this corresponds to a targeted exogenous subsidy to exit that decreases group  $i$ 's exit cost from  $D$  to  $D - W_i$ . This intervention can entail both the stick and the carrot. On the one hand, targeted military operations from international peace-keeping forces may increase the cost of staying in the contest. On the other hand, the promise of impunity to militia commanders or the prospective integration in the political process of the DRC may increase the attractiveness of leaving the contest. We assume the policy treatment to be sufficiently strong to induce the targeted groups to leave, and study which intervention would be most effective in reducing rent dissipation.

The analysis bears a close similarity with the key-player analysis in Ballester *et al.* (2006). In their language, a key player is the agent whose removal triggers the largest reduction in rent dissipation. Proposition 3 in Appendix C establishes formally that in our model the identity of the key player is related to our centrality measure defined in equation (8).<sup>29</sup>

We start with policies targeting single groups. We exclude the FARDC and the DRC police from the set of potential targets (although we retain mutinies), because we do not view removing local government organizations as a policy-relevant option.<sup>30</sup> Table 4 summarizes the results for the 15 groups whose removal yields the largest reduction in rent dissipation at the baseline estimates of column 4 in Table 1. These groups include the most important actors in the conflict. If we exclude the activity of the FARDC, they jointly account for 82% of the total fighting. A complete list of the groups is provided in Table C.1. For each group, we report the number of its enemies and allies, the observed share in total fighting  $x_k^*(G^b) / \sum_{i=1}^n x_i^*(G^b)$ , the reduction in rent dissipation  $(-\Delta \text{RD}_K^{\hat{\beta}, \hat{\gamma}})$  associated with its removal, and a *multiplier* defined as the ratio between the reduction in rent dissipation and the share in total fighting. The multiplier is a useful measure of the impact of the policy weighted by the importance of the group being removed. The fourth and fifth columns are evaluated at the baseline estimates of  $\beta$  and  $\gamma$  of column 4 in Table 1. In the last two columns, we report intervals centered on the baseline estimates with the range of plus and minus one standard deviation. More precisely, we set  $(\hat{\beta}, \hat{\gamma}) \approx (0.085, 0.063)$  and  $(\hat{\beta}, \hat{\gamma}) \approx (0.142, 0.103)$ . This yields a range of variation of the effects as the externality parameters change.

Two findings are noteworthy. First, although there is a high correlation between the observed contribution of each group to total fighting and the reduction in total fighting associated with its removal, the correlation is significantly below unity for the most active groups. For instance, this correlation is 83% in the subsample of the ten most active groups. Second, there is heterogeneity in the multipliers. Rwanda-backed RCD-G, the most active armed group, accounts for less than 9% of the total military activity in the data. Its removal would reduce aggregate fighting by over 15%, with a multiplier of 1.7. Likewise, Uganda-backed RCD-K accounts for 6% of military activity.

<sup>29</sup>More formally, let  $K$  denote a vector comprising a subset of cardinality  $k$  of the  $n$  groups (where  $1 \leq k < n$ ). We denote by  $G^b \setminus \{K\}$  the network after removing the subset  $K$ . The vector of equilibrium fighting efforts is given by equations which are analogous to equation (20) except that the dimension of the system is reduced by  $k$ , the adjacency matrix is  $\mathbf{A}(G^b \setminus \{K\})$ , and the parameters attached to the network structure are replaced by  $\Lambda^{\hat{\beta}, \hat{\gamma}}(G^b \setminus \{K\})$  and  $\Gamma^{\hat{\beta}, \hat{\gamma}}(G^b \setminus \{K\})$ . We compute the rent dissipation before and after the removal of the subgroup  $K$ . The change in rent dissipation equals  $\Delta \text{RD}_K^{\hat{\beta}, \hat{\gamma}} \equiv \text{RD}^{\hat{\beta}, \hat{\gamma}}(G^b \setminus \{K\}) - \text{RD}^{\hat{\beta}, \hat{\gamma}}(G^b)$ , where  $\text{RD}^{\hat{\beta}, \hat{\gamma}}(G^b \setminus \{K\}) \equiv \sum_{i=1}^k x_i^*(G^b \setminus \{K\})$ .

<sup>30</sup>In addition, we consider the Rwandan army as a single entity, namely, we always simultaneously remove the two separate groups associated by ACLED to Rwanda.

Table 4: Welfare effects of removing individual armed groups.

Group	# Enmities	# Allies	Share fight.	$-\Delta RD$	Multipl.	$-\Delta RD$ ( $\pm 1$ SD)	Multipl. ( $\pm 1$ SD)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RCD-G	14	4	0.087	0.151	1.7	[0.125, 0.181]	[1.4, 2.1]
RCD-K	13	5	0.060	0.094	1.6	[0.070, 0.151]	[1.2, 2.5]
Rwanda	17	9	0.053	0.066	1.2	[0.053, 0.109]	[1.0, 2.0]
LRA	6	1	0.041	0.056	1.4	[0.038, 0.115]	[0.9, 2.8]
FDLR	5	6	0.066	0.055	0.8	[0.059, 0.044]	[0.9, 0.7]
Mayi-Mayi	6	7	0.057	0.046	0.8	[0.054, 0.022]	[1.0, 0.4]
Uganda	13	9	0.043	0.043	1.0	[0.038, 0.048]	[0.9, 1.1]
CNDP	3	2	0.043	0.041	0.9	[0.041, 0.040]	[0.9, 0.9]
MLC	7	4	0.031	0.039	1.3	[0.026, 0.074]	[0.8, 2.4]
UPC	5	1	0.022	0.030	1.4	[0.018, 0.057]	[0.8, 2.6]
Lendu Ethnic Mil.	6	3	0.024	0.022	0.9	[0.039, -0.012]	[1.6, 0.5]
Mutiny FARDC	3	2	0.016	0.016	1.0	[0.009, 0.045]	[0.6, 2.8]
Interahamwe	7	5	0.014	0.014	1.0	[0.024, -0.017]	[1.7, 1.2]
ADF	3	4	0.013	0.012	0.9	[0.011, 0.017]	[0.8, 1.3]
FRPI	2	1	0.009	0.010	1.1	[0.003, 0.031]	[0.4, 3.7]

Note: The computation of the counterfactual equilibrium is based on the baseline point estimates of column 4 in Table 1. For each group, we report the number of its enemies and allies (cols. 1-2); the observed share of total fighting involving this group (col. 3); the counterfactual reduction in rent dissipation associated with its removal (col. 4); a multiplier defined as the ratio of col. 4 over col. 3 (col. 5); the reduction in RD and its associated multiplier for a set of parameters equal to the baseline estimates  $\pm 1$  SD (cols. 6-7).

Its removal would reduce fighting by more than 9%, with a multiplier of 1.6. Removing the LRA would reduce rent dissipation by 6%, a larger effect than that from removing more active groups such as the FDLR, the Mayi-Mayi militia, and the CNDP.

In Tables C.8 and C.9, we consider the effect of a simultaneous removal of two and three groups, respectively. For instance, we find that removing Rwanda and its closest ally, the RCD-G, yields a 24% reduction in fighting activity, significantly larger than their 14% contribution to total violence. There is some complementarity in the joint intervention: the effect of their joint removal is 12% larger than the sum of the individual effects. In the case of three groups, the most effective intervention is the removal of Rwanda in combination with the RCD-G and the CNDP (-29.5%, with a multiplier of 1.6), two of Rwanda’s allies.

Finally, consider the effect of targeting selected subsets of armed groups that have particular connections with each other. The upper panel of Table 5 summarizes the results. At the baseline estimates, removing the 29 groups with a foreign affiliation reduces rent dissipation by 27%, in line with their share in total fighting. We show below that the effect of this intervention increases significantly when we allow an endogenous adjustment of the network. Removing the 11 groups involved in the Ituri conflict causes a reduction in rent dissipation of 9%.<sup>31</sup> Removing the 6 groups associated with the Hutu exodus of Rwanda scales down conflict by a mere 9% (lower than the observed activity of these groups). Removing Uganda, Rwanda, and all their associates reduces fighting by 46%, significantly more than the contribution of these groups to conflict in the data. Finally, removing the 16 groups with more than five enemies reduces fighting by 68%. This is a

<sup>31</sup>Ituri is a province of north-eastern DRC that has witnessed a long-lasting conflict between the agriculturalist Lendu and pastoralist Hema ethnic groups. The apex of the conflict was in 1999-2003, although it continues at a lower level until the current days. The groups involved in this conflict for which we have information include: Front for Patriotic Resistance of Ituri, Hema Ethnic Militia, Lendu Ethnic Militia, Nationalist and Integrationist Front, Ngiti Ethnic Militia, Party for the Unity and Safekeeping of Congo’s Integrity, Popular Front for Justice in Congo, Revolutionary Movement of Congo, Union of Congolese Patriots.

Table 5: Welfare effects of removing selected subsets of armed groups.

Set of Groups	# groups	Sh. fight.	$-\Delta RD$	Multiplier	MAD	New enm. & all. (at the median)	Regression coeffs. [enmities,alliances]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>EXOGENOUS NETWORK</b>							
Foreign Groups	29	0.280	0.268	1.0	–	–	–
Ituri	9	0.086	0.094	1.1	–	–	–
Out of Rwanda	6	0.092	0.087	0.9	–	–	–
Rwa&Uga&ass.	10	0.336	0.456	1.4	–	–	–
Large Groups	16	0.802	0.677	0.8	–	–	–
<b>WITH ENDOGENOUS NETWORK RECOMPOSITION</b>							
Foreign Groups	29	0.280	0.412	1.5	0.029	[-11, +8]	[-0.010, +0.008]
Ituri	9	0.086	0.094	1.1	0	[+0, +0]	[-0.003, +0.011]
Out of Rwanda	6	0.092	0.117	1.3	0.031	[0, +2]	[-0.009, +0.011]
Rwa&Uga&ass.	10	0.336	0.332	1.0	0.044	[+3, -6]	[-0.011, +0.011]
Large Groups	16	0.802	0.719	0.9	0.008	[0, +9]	[-0.003, +0.002]

Note: The computation of the counterfactual equilibrium is based on the baseline point estimates of column 4 in Table 1. For each policy experiment, we display the results with an exogenous network (top panel) and the results with an endogenous network recomposition based on 1,000 Monte Carlo simulations (bottom panel). For each experiment, we report the set of removed groups (col. 1); the number of removed groups (col. 2); the observed share of total fighting involving this set of groups (col. 3); the counterfactual reduction (or its median in the bottom panel) in rent dissipation associated with their removal (col. 4); a multiplier defined as the ratio of col. 4 over col. 3 (col. 5); the Median Absolute Deviation in reduction in RD (col. 6); the post-recomposition number of new enmities and alliances at the median Monte Carlo draw (col. 7); the OLS coefficients of enmities and alliances of a regression across Monte Carlo draws of post-recomposition reduction in RD on reduction in RD (exogenous network) and the post-recomposition numbers of new enmities and alliances (col. 8).

large share, though lower than the 80% share of total fighting they account for in the data. In this case, the model predicts some crowding-in of violence from the surviving groups. Overall, these findings confirm the wisdom that the fragmentation in the DRC conflict makes it difficult for international organizations to deliver a single decisive blow.

## 5.2 ARMS EMBARGO

Forcing armed groups out of the contest may be very costly or even politically infeasible. In this section, we study the effect of a less radical policy that increases the marginal cost of fighting for targeted groups without removing them from the contest. We interpret this intervention as targeted sanctions such as an arms embargo, which constrains the stock of arms and ammunitions at the target groups' disposal, or force them to acquire extra equipment at higher prices in the black market. Formally, we increase the fighting cost in equation (1) from  $-x_i$  to  $-(1 + s_i)x_i$ , where  $s_i$  is the policy parameter capturing the size of the intervention. We continue to measure total rent dissipation as the sum of the fighting efforts of all groups, ignoring the additional cost suffered by the armed groups per unit of fighting. A formal analysis of the equilibrium conditions is provided in Appendix C.

Consider policies targeting individual groups. Figure 3 summarizes the results of a tenfold increase in the marginal cost (i.e.,  $s_i = 9$ ). The most significant gains accrue from targeting the two RCD factions, followed by Rwanda and by the LRA. Interestingly, the effects are never large. An embargo on the RCD-G or one on the RCD-K cause, respectively, a 3% and 2% reduction in total fighting. The interventions have a sizeable effect on the group targeted, typically inducing

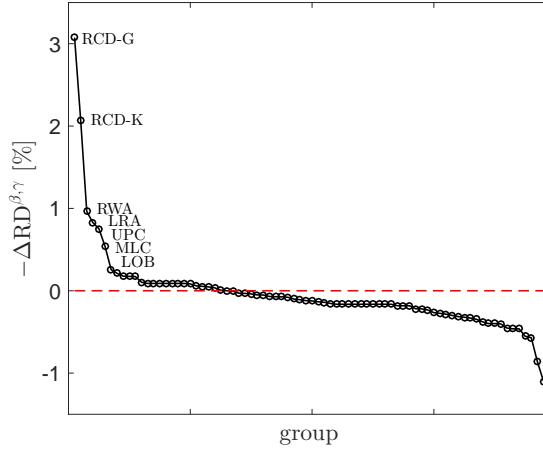


Figure 3: The figure shows the decrease in rent dissipation (relative to the baseline equilibrium) associated with an arms embargo policy targeting each individual group (except the FARDC) separately by setting  $s_i = 9$ . Groups are rank-ordered from the largest to the smallest decrease in rent dissipation. A negative number means that targeting a particular group yields an increase in rent dissipation relative to the benchmark.

a reduction in their fighting activity by 40-60%. However, the non-targeted groups typically fight more, resulting in modest welfare gains. In some cases, embargoes are even counterproductive.

We also consider simultaneously removing multiple groups. Surprisingly, it is never optimal to target many groups. For low levels of  $s_i$ , it is optimal to set an embargo on six groups. However, as we increase  $s_i$ , the cardinality of the optimal number of groups falls. For  $s_i \geq 16$ , it becomes optimal target only a single group.<sup>32</sup>

In summary, the welfare gains of policies that increase fighting costs are small. This result is in line with the skeptical conclusions of studies on the impact of arms embargoes such as Tierney (2005). Recall that in this section we have maintained a prohibitive cost of decommissioning, and only focused on the effect of the policy on an intensive margin. To the extent to which an arms embargo induces a group to drop out of the conflict, the results of Section 5.1 apply.

### 5.3 PACIFICATION POLICIES

In this section, we study the effect of pacification policies aimed at reducing bilateral (or multi-lateral) ethnic and political hostility between groups. More formally, we turn some enmities into neutral links. We view this analysis as especially policy relevant. International organizations may decide to invest in bringing hostile groups to the negotiating table or in de-escalating specific parts of the conflict, subject to limited economic or diplomatic resources (see Hoerner *et al.* 2015). The analysis casts light on which interventions would be most effective.

To provide a benchmark for the potential scope of pacification policies, consider first a drastic counterfactual in which all enmity links are rewired into neutral ones. The effect is large: aggregate fighting is reduced by 65% at the baseline estimates of  $\beta$  and  $\gamma$ . Since wiping out all enmities in the DRC would be utopian, we consider more realistic interventions targeting specific links.

We first consider the effect of pacifying enmity links *vis-a-vis* the FARDC and the DRC police. Table 6 summarizes the results for the 15 groups whose pacification yields the largest reduction in rent dissipation at the baseline estimates. These 15 groups account for 71% of the conflict with the

<sup>32</sup>To see why, consider the case in which  $s_i = 25$  and the RCD-G is subject to an arms embargo. Consider a policy which also targets a second group, the CNDP. Relative to the optimal policy, the fighting effort of the CNDP falls by a fourth. However, the gain is offset by a generalized increase in the fighting of the other groups, and by some bouncing back of the RCD-G effort. Overall, the net effect is more rent dissipation than if the RCD-G were targeted alone. This example is representative of the typical effect of targeting several groups.

Table 6: Welfare effects of pacifying individual armed groups with the FARDC.

Group	Sh. fight.	Sh. bilat. fight.	$-\Delta RD$	Multiplier	$-\Delta RD$ ( $\pm 1$ SD)	Mutipl. ( $\pm 1$ SD)
	(1)	(2)	(3)	(4)	(5)	(6)
Rwanda	0.053	0.010	0.063	6.0	[0.040, 0.140]	[3.8, 13.5]
RCD-G	0.087	0.030	0.056	1.9	[0.033, 0.132]	[1.1, 4.4]
RCD-K	0.060	0.030	0.050	1.6	[0.028, 0.125]	[0.9, 4.1]
LRA	0.041	0.023	0.037	1.6	[0.022, 0.088]	[1.0, 3.9]
MLC	0.031	0.019	0.034	1.8	[0.020, 0.086]	[1.1, 4.6]
Uganda	0.043	0.006	0.031	5.7	[0.020, 0.062]	[3.7, 11.2]
UPC	0.022	0.002	0.024	11.2	[0.014, 0.053]	[6.6, 24.9]
Mutiny FARDC	0.016	0.015	0.023	1.6	[0.015, 0.054]	[1.0, 3.6]
CNDP	0.043	0.038	0.019	0.5	[0.013, 0.033]	[0.4, 0.9]
Lobala Mil.	0.001	0.000	0.017	52.7	[0.011, 0.039]	[32.8, 119.7]
FPJC	0.006	0.006	0.017	2.6	[0.010, 0.037]	[1.7, 5.8]
FRPI	0.009	0.007	0.017	2.4	[0.010, 0.037]	[1.5, 5.3]
BDK	0.002	0.002	0.016	8.1	[0.010, 0.036]	[5.2, 18.3]
Enyele Ethnic Mil.	0.001	0.001	0.016	24.2	[0.010, 0.035]	[15.3, 54.6]
Munzaya Ethnic Mil.	0.001	0.000	0.016	32.3	[0.010, 0.035]	[20.4, 72.7]

Note: The computation of the counterfactual equilibrium is based on the baseline point estimates of column 4 in Table 1. For each group, we report the observed share of total fighting involving this group (col. 1); the observed share of total fighting involving this group against the FARDC (col. 2); the counterfactual reduction in rent dissipation associated with its pacification (col. 3); a multiplier defined as the ratio of col. 3 over col. 2 (col. 4); the reduction in RD and its associated multiplier for a set of parameters equal to the baseline estimates  $\pm 1$  SD (cols. 5-6).

FARDC (and for 70% of the total fighting in the DRC excluding the activity of the FARDC). A complete list is provided in Appendix C.

For each group, we report the observed share in total fighting, the share of total bilateral fighting involving this group and the FARDC, the change in rent dissipation associated with pacifying the link between this group and all factions of the FARDC, and a *multiplier*, defined as the ratio between the third and second columns. Here, the multiplier measures the impact of the policy relative to the size of the conflict between the targeted group and the FARDC. A multiplier of one then means that the pacification yields a mere suppression of the bilateral conflict between two groups. Interestingly, with the exception of the CNDP, all multipliers are well above one, and in some cases are very large. This indicates that pacification induces important spillovers through the network.

The largest absolute gain stems from pacifying the FARDC with Rwanda (6% reduction in fighting), despite the fact that direct military operations between the two armies account for only 1% of total violence. The multiplier of 6 is similar to that of Uganda. Pacifying the FARDC with the two main branches of the RCD is also important. The analysis also identifies a set of small ethnic militias whose pacification with the FARDC would be very effective. In Appendix C, we show the effect of the simultaneous pacification of multiple groups.

We also study the effect of pacifying the same subconflicts discussed in Table 5 above. Here, the policy treatment consists of reconciling all enmities both *vis-a-vis* the FARDC and between the actors in each subconflict. The reconciliation of all foreign groups yields a reduction in rent dissipation of 18%. Interestingly, the reconciliation of the Ituri conflict reduces rent dissipation by 10% – a larger effect than that of wiping out all groups in Table 5. The reconciliation of all groups associated with Uganda and Rwanda yields a 24% reduction in violence.

Finally, we study the effect of pacifying inter- and intra-ethnic conflicts between Hutu- and

Tutsi-affiliated groups.<sup>33</sup> First, we consider rewiring all inter- or intra-Hutu-Tutsi enmities to neutrality. The effect is a reduction in conflict of 9%. The effect becomes much larger if one rewires all bilateral Hutu-Tutsi links to neutrality and all Hutu-Tutsi co-ethnic links to friendships. In this case, the conflict is reduced by 21%.

## 6 ENDOGENOUS NETWORK RECOMPOSITION

In the analysis thus far, we have maintained the assumption of an exogenous network structure, implicitly assuming that alliances and enmities can be traced back to historical relations among groups that are not affected by the warfare dynamics. In some cases (e.g., the historical tensions between Hutus and Tutsis), this is a reasonable assumption. In other cases, such as the alliances forged during the First Congo War, relationships are more malleable. The exogenous network is problematic when we run counterfactual policy experiments. For instance, removing Rwanda or Uganda would likely affect the system of alliances within the DRC.

Ideally, one would like to model a fully endogenous network. There are two main difficulties in our environment. First, for many pairs information is scant, limiting our ability to predict the nature of the link. Second, enmities are by design difficult to rationalize in terms of payoffs, as they often harm both parties involved. Therefore, in a model of endogenous network formation it would be natural to dissolve such links. Since these disadvantageous links exist and persist in the data, they must stem from (often unobservable) historical factors such as grievances over past conflicts.

In this section, we construct a model of semi-endogenous network formation that predicts the resilience of network links to exogenous policy shocks. We postulate a discrete choice Random Utility Model (RUM) where each pair of groups selects the bilateral link (either enmity, alliance, or neutrality) in order to maximize utility. We make the strong assumption that the formation of the link  $\{i, j\}$  depends on the characteristics of  $i$  and  $j$ , including their position in the network, being otherwise independent of all other links. The assumption of conditional independence could be violated in microfounded models where each group rationally decides which links to add or break, and spillovers across different decisions arise.

Our approach is close in spirit to Fafchamps and Gubert (2007), although in their papers interactions between groups have a binary nature. It is also close to Leskovec *et al.* (2010), who use a logistic regression to estimate signed networks, and to Jiang (2015), who studies a stochastic block model for signed graphs.

### 6.1 RANDOM UTILITY MODEL

We estimate a choice model of link formation that is based on the following RUM:

$$U_{ij}(a) = \alpha \times CSF_{ij}(a) + \mathbf{X}_{ij} \times \boldsymbol{\xi}(a) + \mathbf{Z}_{ij} \times \boldsymbol{\zeta}(a) + FE_i(a) + FE_j(a) + \tilde{u}_{ij}(a), \quad (21)$$

where  $a \in \{-1, 0, 1\}$ , and  $U_{ij}$  is the joint utility of dyad  $ij$  associated with the alternative  $a$ . Each dyad chooses the link that maximizes its surplus,  $a_{ij}^* = \arg \max U_{ij}(a)$ . We abstract from distributional issues by assuming that each dyad makes the efficient choice and can then arrange within-dyad transfers so as to ensure that the choice is acceptable to both parties. The utility of each of the three alternatives depends on observable and unobservable factors comprising:

1.  $CSF_{ij}(a)$ , the equilibrium joint payoff of the dyad  $ij$  in the second stage CSF game (equation 1) where the network structure has alternative  $a$  for link  $\{i, j\}$ , the other network links being

---

<sup>33</sup>We code Tutsi and Hutu affiliation following the ethnic group data from Cederman *et al.* (2009). In our sample, 14 groups are ethnical Tutsis, and 11 are ethnical Hutus.



unchanged.  $CSF_{ij}$  can be inferred from our structural equation (20) once the parameters  $\beta$  and  $\gamma$  are known and the network structure  $G$  is adjusted for alternative  $a$ .

2.  $\mathbf{X}_{ij}$ , a vector of dyad-specific characteristics including the spatial distance between the centroids of  $i$  and  $j$ , and categorical variables capturing the fact that they are affiliated to the same ethnic group (from Cederman *et al.* 2009), and whether they have a common or opposite Tutsi-Hutu background, whether at least one of them is a foreign army, whether at least one of them is a government actor. These characteristics are likely predictors of patterns of alliance or enmity. For instance, the Hutu-Tutsi antagonism is expected to increase the utility associated with  $a = -1$ .
3.  $\mathbf{Z}_{ij}$ , network-dependent characteristics, in the spirit of Leskovec *et al.* (2010) that are likely to have a systematic effect on the nature of the link. These comprise the number of common allies and common enemies of  $i$  and  $j$ , and the number of common conflicting neighbors (namely,  $i$ 's enemies that are  $j$ 's allies, or vice versa).
4. Alternative-dependent group fixed effects  $FE_i(a)$  that capture the unconditional propensity of  $i$  to form the alternative  $a$ .
5.  $\tilde{u}_{ij}(a)$ , random utility shocks assumed to have a type I extreme value distribution.

This model is estimated by maximum likelihood as a standard conditional logit estimator. We run an alternative-specific conditional logit for  $a_{ij} = +1$  (alliance) and  $a_{ij} = -1$  (enmity), setting neutrality as the reference state. This yields an estimated probability that the link  $\{i, j\}$  is an alliance, resp. an enmity, relative to neutrality. The estimation results are reported in Appendix Table B.12.

The coefficient  $\alpha$  – the only coefficient that is not alternative-specific – is insignificant, implying that conflict-specific payoffs under the different alternatives have very low predictive power. This finding is reassuring, being consistent with our assumption that the network structure is exogenous to our baseline CSF game. In contrast, both  $\mathbf{X}_{ij}$  and  $\mathbf{Z}_{ij}$  have significant explanatory power, with signs broadly in line with prior expectations. In particular, if  $i$  is Hutu and  $j$  is Tutsi (or vice versa) the probability that  $i$  and  $j$  are enemies is significantly higher than that they are allies or have a neutral relationship. As expected, spatial proximity is a strong positive predictor of both alliances and enmities. Moving to network-dependent characteristics, if  $i$  and  $j$  have common enemies, it is more likely that they are allies and less likely that they are enemies (both effects being highly significant). Similarly, if they have relationships of opposite sign with a third group (e.g.,  $i$  is an enemy of  $k$ , while  $j$  is an ally of  $k$ ), then,  $i$  and  $j$  are more likely to be enemies and less likely to be allies. More surprisingly, common alliances decrease the probability for the two groups to be allied – the effect on being enemies being close to zero. This is in line with the narrative that many links are non-transitive. The model fits the data well. Appendix B provides details about the goodness of fit.

## 6.2 RE-ESTIMATING THE MODEL USING THE NETWORK STRUCTURE PREDICTED BY THE RUM

The main goal is to use the RUM to predict the changes in the network structure induced by policy shocks. Before turning to that, we take a brief de-tour to re-estimate the model using the network structure predicted by the RUM as an instrument for the observed network. Consider the regression equation (15), and in particular the IV regression where  $TFE_{it}$  and  $TFA_{it}$  are instrumented by the rainfall in allies' and enemies' territories. Here, we replace  $RE_{it}$  and  $RA_{it}$  by

$\widehat{RE}_{it} = \sum_{j=1}^n \widehat{p}_{ij}^- \times \text{RAIN}_{jt}$  and  $\widehat{RA}_{it} = \sum_{j=1}^n \widehat{p}_{ij}^+ \times \text{RAIN}_{jt}$ , where  $\widehat{p}_{ij}^- \in [0, 1]$  and  $\widehat{p}_{ij}^+ \in [0, 1]$  are the probabilities that groups  $i$  and  $j$  are allies and enemies as predicted by the RUM.<sup>34</sup> Relative to the baseline estimation,  $\widehat{p}_{ij}^-$  and  $\widehat{p}_{ij}^+$  replace the (observed) links,  $a_{ij}^-$  and  $a_{ij}^+$ , in the construction of the instruments for the 2SLS estimator. Thus, the exogenous source of variation is the rainfall shocks in other groups' territories and the set of dyadic characteristics (ethnicity, spatial proximity, etc.) and network-specific covariates in equation (21). Note that the results are robust to restricting the RUM to the set of dyadic characteristics only.

The results are shown in column 9 of Table B.3, which displays the analogue of the baseline estimation in column 4 of Table 1.<sup>35</sup> The coefficients are in the ball park of the baseline estimates. The coefficients of TFE and TFA are, respectively, 0.11 (s.e. 0.03) and -0.11 (s.e. 0.05). In spite of the low KP-stat of 4 indicating a weak instrument problem, we find the results reassuring, given the challenge of estimating a complex network like the one in the DRC war.

### 6.3 ENDOGENOUS NETWORK ADJUSTMENTS AFTER POLICY SHOCKS

In this section, we use our estimated choice model of link formation to predict the changes in the network structure triggered by policy shocks. This intervention affects both  $CSF_{ij}(a)$  and  $\mathbf{Z}_{ij}$  in equation (21), which in turn affects the prediction of the RUM. We allow for post-intervention network recomposition and quantify the impact of the policy on fighting in the recomposed network.<sup>36</sup>

Since the conditional logit model does not yield estimates of the unobserved random utility shocks  $\tilde{u}_{ij}(a)$  in equation (21), our analysis must rely on Monte Carlo simulations. More precisely, for each policy experiment, we perform Monte Carlo simulations of the network recomposition and obtain a counterfactual distribution of fighting efforts. In the tables below, we focus on the effects at the median realization, although in some cases we show the entire distribution.

For a given policy experiment, we iterate the following algorithm 1,000 times:

1. We draw a vector of random utility shocks  $\tilde{\mathbf{u}}_{ij}$  for each dyad  $ij$  from a truncated multivariate type I extreme value distribution with unconditional mean and variance being, respectively, 0.577 (the Euler-Mascheroni constant) and  $\sqrt{\pi}/6$ . The support of the distribution corresponds to the domain of  $\tilde{\mathbf{u}}_{ij}$  that is compatible with the link observed in the data. This ensures that, in the absence of policy intervention, there is no network recomposition.<sup>37</sup>

<sup>34</sup>The prediction of the observed component of utility in equation (21) is given by

$$\widehat{V}_{ij}(a|G, \mathbf{X}_{ij}, \mathbf{Z}_{ij}) = \hat{\alpha} \times CSF_{ij}(a) + \mathbf{X}_{ij} \times \widehat{\boldsymbol{\xi}}(a) + \mathbf{Z}_{ij} \times \widehat{\boldsymbol{\zeta}}(a) + \widehat{FE}_i(a) + \widehat{FE}_j(a)$$

with the normalization  $\boldsymbol{\xi}(0) = \boldsymbol{\zeta}(0) = FE_i(0) = FE_j(0) = 0$ . In turn, the predicted conditional probabilities are given by the standard formula  $\mathbb{P}_{ij}(a|G, \mathbf{X}_{ij}, \mathbf{Z}_{ij}) = e^{V_{ij}(a)} / (e^{V_{ij}(-1)} + e^{V_{ij}(0)} + e^{V_{ij}(+1)})$ . Henceforth,  $\widehat{p}_{ij}^- \equiv \mathbb{P}_{ij}(a = -1|G, \mathbf{X}_{ij}, \mathbf{Z}_{ij})$  and  $\widehat{p}_{ij}^+ \equiv \mathbb{P}_{ij}(a = +1|G, \mathbf{X}_{ij}, \mathbf{Z}_{ij})$ .

<sup>35</sup>The results are also robust when we consider a time-varying network as in Section 4.2.1. Appendix Table B.13 shows the set of main results for the benchmark specifications in Tables 1 and 3.

<sup>36</sup>Note that removing groups and pacification policies affect both  $CSF$  and  $\mathbf{Z}$ . In contrast, embargoes only affect  $CSF$ . Since, as we show below, the estimated effect of  $CSF$  is small, embargoes do not trigger network recomposition in our model.

<sup>37</sup>Our sampling procedure for drawing the random utility shocks from a truncated multivariate distribution follows a standard accept-reject algorithm (see Train 2003, Chapter 9), where the definition of the acceptance domain  $\tilde{\mathbf{u}}_{ij}$  follows from the RUM. We first draw a candidate triplet  $\tilde{\mathbf{u}}_{ij}$  from the unconditional density. Denoting by  $V_{ij}(a)$  the observed utility in equation (21), we retain the draw if it is compatible with the observed link  $a_{ij}^{obs}$ , namely  $V_{ij}(a_{ij}^{obs}) + \tilde{u}_{ij}(a_{ij}^{obs}) = \max_{a \in \{-1, 0, +1\}} V_{ij}(a) + \tilde{u}_{ij}(a)$ . If the triplet does not satisfy the previous condition, we reject this draw and we draw a new triplet. The procedure stops when 1,000 accepted draws have been obtained for each dyad  $ij$ .

This conditional approach implies that the simulated network recompositions are entirely driven by the policy-

2. For each dyad  $ij$ , we compute the post-policy values of  $CSF_{ij}$  and  $\mathbf{Z}_{ij}$  (the other covariates in equation (21) are not affected). Given these and the estimated parameters, we compute the after-policy observable component of utility  $V_{ij}^{post}(a)$ .
3. For each Monte Carlo draw, we compute the post-policy optimal link:  $a_{ij}^{post} = \arg \max_{a \in \{-1, 0, +1\}} V_{ij}^{post}(a) + \tilde{u}_{ij}(a)$ . Rewiring occurs when  $a_{ij}^{obs} \neq a_{ij}^{post}$ . This yields the post-policy network  $G^{post}$ .
4. The counterfactual equilibrium vector of fighting efforts is obtained from the structural equation (20) and  $G^{post}$ .

Typically, allowing for network recomposition in response to policy shocks increases (decreases) rent dissipation relative to the exogenous network benchmark whenever the policy shock triggers an increase (reduction) in the number of enmities and a reduction (increase) in the number of alliances. Note that a policy may affect both the number of alliances and enmities in the same direction, causing ambiguous net effects.

### 6.3.1 Removing Armed Groups

In this section, we study the effect of removing one or more groups from the conflict (cf. Section 5.1) when endogenous network adjustments are allowed. Consider Figure 4 (more details can be found in Appendix Table B.14). The left panel shows the reduction in rent dissipation with and without network recomposition for the top 15 groups (all but three groups remain in the top 15 even after allowing for network recomposition). The correlation is high (81%), implying that the short-run effects of Section 5.1 are overall robust to network recomposition. Among the groups whose removal causes the largest network recomposition, we find the armies of Rwanda and Uganda. Recall that removing Rwanda causes a reduction in rent dissipation of 6.6% when the network is exogenous. The adjustment of the network causes a further 3.7% reduction, lifting the median total effect of removing Rwanda to 10.3% (more than twice as large as its observed fighting share). This additional effect is due to five enmities switching to neutral links. Similarly, removing Uganda triggers some network recomposition (four enmities destroyed, four alliances formed). Two groups whose removal is especially consequential for the network structure are the Lendu Ethnic Militia and the ADF. In both cases, the indirect effect of removing them from the contest exceeds the direct effect of the policy under an exogenous network.

Consider, next, the effect of removing selected subgroups of armed groups. The results are reported in the lower panel of Table 5. The most remarkable new result is in the experiment where we remove all groups with a foreign affiliation. In this case, the reduction in rent dissipation increases from 27% (exogenous network) to 41%. The effect is estimated precisely, with a median absolute deviation (MAD) of 4.1%. The large extra reduction in fighting efforts accrue from both a reduction in the number of enmities (nine at the median) and an increase in the number of friendly links (six at the median).

The effect of removing the groups associated with the Hutu exodus is also magnified significantly by the network recomposition (from 8.7% to 12.2%). The same is true for the set of large groups. In other cases, the recomposition of the network has an attenuating effect or no effect.

---

driven changes in  $CSF_{ij}(a)$  and  $\mathbf{Z}_{ij}$  and not by re-sampling of the unobserved utility shocks.

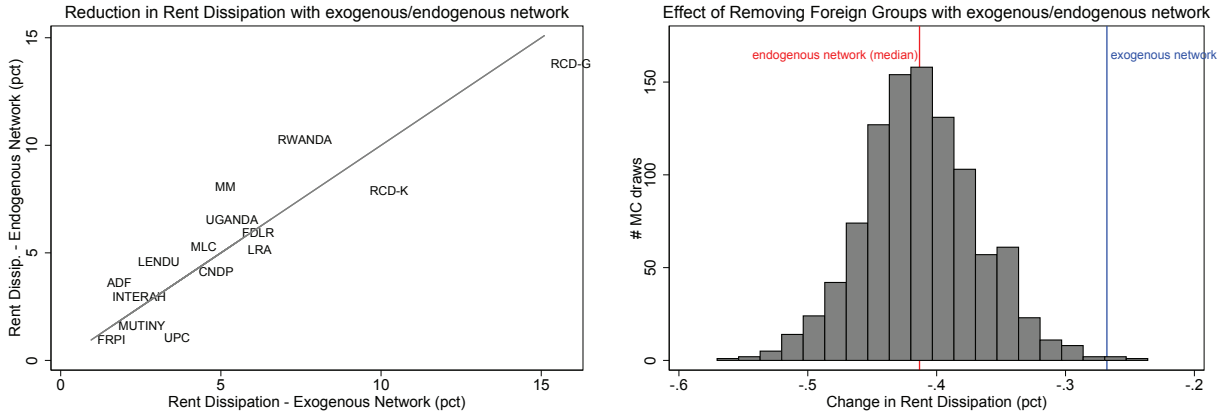


Figure 4: (Left panel) This figure displays the reduction in rent dissipation with and without network recomposition following the removal of each of the top 15 groups. (Right panel) This figure displays the Monte Carlo distribution of the reduction in rent dissipation following the removal of 24 foreign groups (1,000 simulated network recompositions).

## 7 CONCLUSION

In this paper, we construct a theory of conflict in which different groups compete over a fixed amount of resources. We introduce a network of alliances and enmities that we model as externalities in a Tullock contest success function. Alliances are beneficial to each member, but are not unitary coalitions. Rather, each group acts strategically *vis-a-vis* both allies and enemies. We view our theory as especially useful in conflicts characterized by high fragmentation, non-transitive relations, and decentralized military commands, all common features of civil conflicts.

We apply the theory to the analysis of the Second Congo War. Our estimation of the network externalities tackles a reflection problem through an instrumental variable strategy. The signs of the estimated coefficients conform with the prediction of the theory. Each group’s fighting effort is increasing in the total fighting of its enemies and decreasing in the total fighting of its allies. We then use our structural model to quantify the effect of different pacification policies. In an extension, we allow not only the groups’ fighting effort but also the network of alliances and enmities to respond endogenously to policy interventions.

Informal alliances and enmities and intransitive links are not unique to the Congo. Rather, they are common in many contemporary conflicts such as Afghanistan, Somalia, Iraq, Sudan, and Syria. While the results of the various policy experiments cannot be mechanically generalized to other conflicts, our analysis provides new methodological tools to study how webs of alliances and enmities can lead to escalation or containment of conflict. Future work can build on this to propose a full-fledged model of endogenous network formation. In work in progress, we are extending the analysis to other fragmented conflicts such as the recent civil war in Syria.

## REFERENCES

- ACEMOGLU, D., C. GARCIA-JIMENO, AND J. ROBINSON (2015): “State Capacity and Economic Development: A Network Approach,” *American Economic Review*, **105**(8), 2364–2409.
- AUTESSERRE, S. (2008): “The Trouble with Congo: How Local Disputes Fuel Regional Conflict,” *Foreign Affairs*, **87**, 94–110.
- BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2006): “Who’s who in networks. Wanted: The key player,” *Econometrica*, **74**(5), 1403–1417.
- BATES, R., A. GREIF, AND S. SINGH (2002): “Organizing Violence,” *Journal of Conflict Resolution*, **46**(5), 599–628.

- BERMAN, N., M. COUTTENIER, D. ROHNER, AND M. THOENIG (2014): “This mine is mine! How minerals fuel conflicts in Africa,” *CEPR Discussion Paper 10089*.
- BLOCH, F. (2012): “Endogenous formation of alliances in conflicts,” in *The Oxford Handbook of the Economics of Peace and Conflict*, ed. by Michelle Garfinkel, and Stergios Skaperdas. Oxford University Press.
- BLUNDELL, R., AND J.M. ROBIN (1999): “Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems,” *Journal of Applied Econometrics*, **14**(3), 209-232.
- BRAMOULLÉ, Y., H. DJEBBARI, AND B. FORTIN (2009): “Identification of peer effects through social networks,” *Journal of Econometrics*, **150**(1), 41-55.
- BRAMOULLÉ, Y., R. KRANTON, AND M. D’AMOURS (2014): “Strategic Interaction and Networks,” *American Economic Review*, **104**(3), 898-930.
- BRÜCKNER, M., AND A. CICCONE (2011): “Rain and the democratic window of opportunity,” *Econometrica*, **79**(3), 923-47.
- CASSAR, A., P. GROSJEAN, AND S. WHITT (2013): “Legacies of violence: trust and market development,” *Journal of Economic Growth*, **18**(3), 285-318.
- CEDERMAN, L.E., H. BUHAUG, AND J.K. ROD (2009): “Ethno-Nationalist Dyads and Civil War: A GIS-Based Analysis,” *Journal of Conflict Resolution*, **53**(4), 496-525.
- CVETKOVIC, D., M. DOON, AND H. SACHS (1995): *Spectra of Graphs: Theory and Applications*. Leipzig: Johann Ambrosius Barth.
- CHANDRASEKHAR, A., AND R. LEWIS (2016): “Econometrics of Sampled Networks,” *mimeo*, Stanford.
- CONLEY, T.G. (1999): “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, **92**(1), 1-45.
- CONLEY, T.G. (2008): “Spatial Econometrics,” in *The New Palgrave Dictionary of Economics*, ed. by Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.
- CUNNINGHAM, D., K. GLEDITSCH, AND I. SALEHYAN (2013): “Non-State Actor Data,” Dataset (Version 3.4, 23/11/2013), <http://privatwww.essex.ac.uk/~char126\relaxksg/eacd.html>.
- DELL, M. (2012): “Path dependence in development: Evidence from the Mexican Revolution,” *mimeo*, Harvard
- DUBE, O., AND J. VARGAS (2013): “Commodity Price Shocks and Civil Conflict: Evidence from Colombia,” *Review of Economics Studies*, **80**(4), 1384-1421.
- ECK, K. (2012): “In data we trust? A comparison of UCDP GED and ACLED conflict events datasets,” *Cooperation and Conflict*, **47**(1), 124-141.
- ESTEBAN, JOAN AND, DEBRAJ RAY (2001): “Collective Action and the Group Size Paradox,” *American Political Science Review*, **95**(3), 663-672.
- FAFCHAMPS, M., AND F. GUBERT (2007): “Risk Sharing and Network Formation,” *American Economic Review*, **97**(2), 75-79.
- FEARON, J.D. (2008): “Economic Development, Insurgency, and Civil War,” in *Institutions and Economic Performance*, ed. by Elhanan Helpman. Harvard University Press.
- FRANKE, J., AND T. ÖZTÜRK (2015): “Conflict networks,” *Journal of Public Economics*, **126**, 104-113.
- GROSSMAN, H., AND M. KIM (1995): “Swords or Plowshares? A Theory of the Security of Claims to Property,” *Journal of Political Economy*, **103**(6), 1275-1288.
- HIDALGO, F.D., S. NAIDU, S. NICTER, AND N. RICHARDSON (2010): “Economic Determinants of Land Invasions,” *Review of Economics and Statistics*, **92**(3), 505-523.
- HILLER, T. (2017): “Friends and Enemies: A Model of Signed Network Formation,” *forthc.* in *Theoretical Economics*.
- HIRSHLEIFER, J. (1989): “Conflict and rent-seeking success functions: Ratio vs. difference models of relative success,” *Public Choice*, **63**(2), 101-112.
- HOERNER, J., M. MORELLI, AND F. SQUINTANI (2015): “Mediation and Peace,” *Review of Economic Studies*, **82**(4), 1483-1501.
- HSIANG, S., M. BURKE, AND E. MIGUEL (2013): “Quantifying the Influence of Climate on Human Conflict,” *Science*, **341**(6151).

- HUREMOVIC, K. (2014): "Rent Seeking and Power Hierarchies: A Noncooperative Model of Network Formation with Antagonistic Links", *Nota di Lavoro*. Fondazione Eni Enrico Mattei.
- INTERNATIONAL CRISIS GROUP (1998): "Congo at War: A Briefing on the Internal and External Players in the Central African Conflict," <http://www.crisisgroup.org/>.
- JACKSON, M. (2008): *Social and Economic Networks*. Princeton University Press.
- JACKSON, M., AND S. NEI (2014): "Networks of Military Alliances, Wars, and International Trade," *Proceedings of the National Academy of Sciences*, **112**(50), 15277-15284.
- JACKSON, M., AND Y. ZENOU (2014): "Games on networks," in *Handbook of Game Theory*, ed. by P. Young and S. Zamir, vol. 4. Elsevier, Amsterdam.
- JEANTY, P.W. (2012): "Stata Implementation of the Non-Parametric Spatial Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *mimeo*, Rice University.
- JIANG, J. (2015): "Stochastic block model and exploratory analysis in signed networks," *Physical Review*, **E91**(6), 1-12.
- KLEINBERGEN, F., AND R. PAAP (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, **133**(1), 97-126.
- KONRAD, K. (2011): "Aspects of Fighting in Alliances," *mimeo*, Max Planck Inst. for Tax Law and Public Fin.
- HARARI, M.F., AND E. LA FERRARA (2012): "Conflict, Climate and Cells: A Disaggregated Analysis," *IGIER Working Paper 461*.
- LESKOVEC, J., D. HUTTENLOCHER, AND J. KLEINBERG (2010): "Predicting Positive and Negative Links in Online Social Networks," *WWW'10 Proceedings of the 19th international conference on World wide web*, 641-650.
- McKNIGHT, S. (2015): "The Rise and Fall of the Rwanda-Uganda Alliance (1981-1999)," *African Studies Quarterly*, **15**(2), 23-52.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2013): "The Long-Run effects of the Scramble for Africa," *NBER Working Paper 17620*.
- MIGUEL, E., S. SATYANATH, AND E. SERGENTI (2004): "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, **112**(4), 725-753.
- MURDOCK, G.P. (1959): *Africa: Its Peoples and Their Culture History*. McGraw-Hill, New York.
- NUNN, N., AND L. WANTCHEKON (2011): "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, **101**(7), 3221-3252.
- OLSON, M., AND R. ZECKHAUSER (1966): "An economic theory of alliances," *Review of Economics and Statistics*, **48**(3), 266-279.
- PRUNIER, G. (2011): *Africa's World War: Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford University Press, Oxford.
- ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2013a): "War Signals: A Theory of Trade, Trust and Conflict," *Review of Economic Studies*, **80**(3), 1114-1147.
- ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2013b): "Seeds of Distrust? Conflict in Uganda," *Journal of Economic Growth*, **18**(3), 217-252.
- SANCHEZ DE LA SIERRA, R. (2016): "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo," *mimeo*, Columbia University.
- SCHNEIDER, U., A. BECKER, P. FINGER, A. MEYER-CHRISTOFFER, B. RUDOLF, AND M. ZIESE (2011): "GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data," DOI: 10.5676/DWD\_GPCC/FD\_M\_V6\_050.
- SEYBOLT, T. (2000): "Appendix 1B: The war in the Democratic Republic of Congo," in *SIPRI Yearbook 2000*, ed. by Stockholm International Peace Research Institute. Oxford University Press, Oxford.
- SKAPERDAS, S. (1996): "Contest Success Functions," *Economic Theory*, **7**(2), 283-290.
- STEARNS, J. (2011): *Dancing in the Glory of Monsters: The Collapse of the Congo and the Great War of Africa*. PublicAffairs, New York.

- SUNDBERG, R., AND E. MELANDER (2013): “Introducing the UCDP Georeferenced Event Dataset,” *Journal of Peace Research*, **50**(4), 523-532.
- TIERNEY, D. (2005): “Irrelevant or malevolent? UN arms embargoes in civil wars,” *Review of International Studies*, **31**(4), 645-664.
- TRAIN, K.E. (2009): *Discrete Choice Methods with Simulation*. Cambridge University Press.
- TULLOCK, G. (1980): “Efficient Rent Seeking,” in “*Toward a theory of the rent-seeking society*,” ed. by James Buchanan, Robert Tollison, and Gordon Tullock. Texas A&M University Press, Texas.
- TURNER, T. (2007): *The Congo Wars: Conflict, Myth and Reality*. Zed Books, London.
- VANDEN EYNDE, O. (2017): “Targets of Violence: Evidence from India’s Naxalite Conflict,” forthc. in *Economic Journal*.
- VAN DER WINDT, P., AND M. HUMPHREYS (2016): “Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time,” *Journal of Conflict Resolution*, **60**(4), 748-781
- VOGELSANG, T. J. (2011): “Heteroskedasticity, Autocorrelation, and Spatial Correlation Robust Inference in Linear Panel Models with Fixed-Effects,” *mimeo*, Michigan State University.
- WHITE, S. (2014): “Now What? The International Response to Internal Displacement in the DRC,” Brookings Institute.
- WILLIAMS, C. (2013): “Explaining the Great War in Africa: How Conflict in the Congo Became a Continental Crisis,” *Fletcher Forum of World Affairs*, **37**, 81-100.
- ZEENDER, G., AND J. ROTHING (2010): “Displacement trends in DRC,” *Forced Migration Rev.* **36**, 10-12.

## A THEORY APPENDIX

## A.1 PROOF OF PROPOSITION 1

We first establish the existence of a Nash Equilibrium in which all groups participate in the contest (an interior equilibrium). Let  $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^\top \in \mathbb{R}^n$  denote the candidate equilibrium effort vector that satisfies the FOCs; let  $\mathbf{x}_{-i}^* \in \mathbb{R}^{n-1}$  denote the same vector without the  $i$ 'th component. Let  $\pi_i(G; x_i, \mathbf{x}_{-i}^*) = \varphi_i(G; x_i, \mathbf{x}_{-i}^*) / \sum_{j=1}^n \varphi_j(G; x_i, \mathbf{x}_{-i}^*) - x_i$  denote the payoff function of a deviation from the equilibrium effort, in the range where  $\varphi_i \geq 0$ .

The FOCs of the profit maximization problem yields:

$$0 = \frac{\partial \pi_i(G; x_i^*, \mathbf{x}_{-i}^*)}{\partial x_i} = \frac{\sum_{j=1}^n \varphi_j^* - \varphi_i^*(1 + \beta d_i^+ - \gamma d_i^-)}{\left(\sum_{j=1}^n \varphi_j^*\right)^2} - 1. \quad (22)$$

Here we have used the fact that  $\partial \varphi_j^*(G; x_i, \mathbf{x}_{-i}^*) / \partial x_i = \delta_{ij} + \beta a_{ij}^+ - \gamma a_{ij}^-$  (where  $\delta_{ij} = 0$  if  $i \neq j$  and  $\delta_{ii} = 1$ ), consequently,  $\sum_{j=1}^n \partial \varphi_j^* / \partial x_i = 1 + \beta d_i^+ - \gamma d_i^-$ . Standard algebra yields:

$$\varphi_i^* = \frac{1}{1 + \beta d_i^+ - \gamma d_i^-} \left(1 - \sum_{j=1}^n \varphi_j^*\right) \sum_{j=1}^n \varphi_j^*. \quad (23)$$

Next, define  $\Gamma_i^{\beta, \gamma}(G) \equiv (1 + \beta d_i^+ - \gamma d_i^-)^{-1} > 0$  and  $\Lambda^{\beta, \gamma}(G) \equiv 1 - \left(\sum_{i=1}^n \Gamma_i^{\beta, \gamma}(G)\right)^{-1}$ , where the inequality follows from (3). Summing over  $i$ 's in equation (23) implies that

$$\varphi_i^* = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G))\Gamma_i^{\beta, \gamma}(G) > 0. \quad (24)$$

The inequality hinges on establishing that  $\Lambda^{\beta, \gamma}(G) > 0$ , or equivalently  $\sum_{i=1}^n \Gamma_i^{\beta, \gamma}(G) > 1$ . Observe that  $\sum_{i=1}^n \Gamma_i^{\beta, \gamma}(G) = \sum_{i=1}^n \frac{1}{1 + \beta d_i^+ - \gamma d_i^-} \geq \sum_{i=1}^n \frac{1}{1 + \beta d_i^+} \geq \frac{n}{1 + \beta d_{\max}^+} > 1$ . The last inequality holds true if and only if  $\beta < \frac{n-1}{d_{\max}^+}$ , which is in turn is necessarily true if  $\beta < 1$ . This, in turn, follows from the assumption that  $\beta + \gamma < 1 / \max\{\lambda_{\max}(G^+), d_{\max}^-\}$  that implies that  $\beta + \gamma < 1 / \max\{\lambda_{\max}(G^+), \lambda_{\max}(G^-)\}$ , since  $\lambda_{\max}(G^-) < d_{\max}^-$  (Cvetkovic et al. 1995). Moreover, for any non-empty graph  $G$ ,  $\lambda_{\max}(G) \geq 1$ , because for any graph  $G$ ,  $\lambda_{\max}(G) \geq \max_{i=1, \dots, n} \sqrt{d_i}$  (Cvetkovic et al. 1995), and  $\max_{i=1, \dots, n} d_i \geq 1$  when  $G$  is not empty. Thus,  $\beta \leq \beta + \gamma < 1$ . This establishes that  $\varphi_i^* \geq 0$  for all  $i = 1, \dots, n$ .

Next, we compute  $\mathbf{x}^*$ . Combining (2) with (24) yields:

$$x_i^* + \beta \sum_{j=1}^n a_{ij}^+ x_j^* - \gamma \sum_{j=1}^n a_{ij}^- x_j^* = \Lambda^{n, \beta}(G)(1 - \Lambda^{n, \beta}(G))\Gamma_i^{\beta, \gamma}(G). \quad (25)$$

Denoting by  $\mathbf{\Gamma}^{\beta, \gamma}(G) \equiv (\Gamma_1^{\beta, \gamma}(G), \dots, \Gamma_n^{\beta, \gamma}(G))^\top$ , we can write this system in matrix form as

$$(\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-) \mathbf{x}^* = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G)) \mathbf{\Gamma}^{\beta, \gamma}(G). \quad (26)$$

The fact that  $\beta + \gamma < 1 / \max\{\lambda_{\max}(G^+), \lambda_{\max}(G^-)\}$  also ensures that the matrix  $\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-$  is invertible.<sup>38</sup> Then, (26) yields the effort levels:

$$\mathbf{x}^* = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G)) \mathbf{c}^{\beta, \gamma}(G), \quad (27)$$

<sup>38</sup>This follows from Weyl's Theorem. The determinant of a matrix of the form  $\mathbf{I}_n - \sum_{j=1}^p \alpha_j \mathbf{W}_j$  is strictly positive if  $\sum_{j=1}^p |\alpha_j| < 1 / \max_{j=1, \dots, p} \|\mathbf{W}_j\|$ , where  $\|\mathbf{W}_j\|$  is any matrix norm, including the spectral norm, which corresponds to the largest eigenvalue of  $\mathbf{W}_j$ .



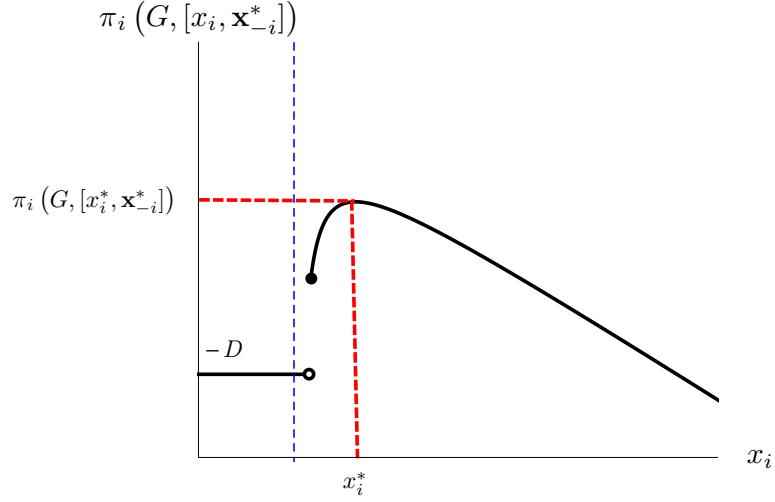


Figure A.1: The figure shows the function  $\pi_i(G, [x_i, \mathbf{x}_{-i}^*])$  for different values of  $x_i$ .

where  $\mathbf{c}^{\beta, \gamma}(G)$  is the centrality measure defined by equation (8)). Equation (27) is the matrix-form version of equation (7) in the proposition. Evaluating  $\pi_i(G, \mathbf{x})$  at  $\mathbf{x} = \mathbf{x}^*$  yields equation (9) in the proposition.

Thus far, we have established that  $\mathbf{x}^*$  and  $\varphi^*$  satisfy the FOCs. In order to prove that the FOCs pin down a Nash equilibrium, we must establish that, for all  $i = 1, 2, \dots, n$ ,  $x_i^*$  is a global maximum of  $\pi_i(G; x_i, \mathbf{x}_{-i}^*)$  for all  $x_i \in \mathbb{R}$ . To prove the result, we split the horizontal line at the cutoff value  $\hat{x}_i$ , uniquely defined by the condition  $\varphi_i(G, \hat{x}_i, \mathbf{x}_{-i}^*) = 0$ . For  $x_i < \hat{x}_i$ ,  $\pi_i(G; x_i, \mathbf{x}_{-i}^*) = -D$ . For  $x_i \geq \hat{x}_i$ , standard algebra establishes that  $(\partial^2 \pi_i / \partial x_i^2)(G; x_i, \mathbf{x}_{-i}^*) = -2 / (\Gamma_i^{\beta, \gamma}(G) \times \Lambda^{\beta, \gamma}(G)) < 0$ , where the inequality follows from the facts, established above, that  $\Gamma_i^{\beta, \gamma}(G) > 0$  and  $\Lambda^{\beta, \gamma}(G) > 0$ . Thus,  $\pi_i(G, x_i, \mathbf{x}_{-i}^*)$  is strictly concave in  $x_i$  in the subdomain  $x_i \geq \hat{x}_i$ . Moreover, equation (4) establishes that  $\varphi_i^* > 0 = \varphi_i(G, \hat{x}_i, \mathbf{x}_{-i}^*)$ . This, together with the fact that  $\varphi_i$  is increasing in  $x_i$ , establishes that  $x_i^* > \hat{x}_i$ . The facts that (i)  $\pi_i(G, x_i, \mathbf{x}_{-i}^*)$  is strictly concave in  $x_i$ , and (ii)  $x_i^* > \hat{x}_i$  jointly imply that  $\pi_i(G, x_i^*, \mathbf{x}_{-i}^*)$  is a global maximum of the  $\pi_i$  function in the subdomain  $x_i \geq \hat{x}_i$ . It is immediate that  $\pi_i(G, x_i^*, \mathbf{x}_{-i}^*) < \infty$ . Define  $\underline{D} = \max_i -\pi_i(G, x_i^*, \mathbf{x}_{-i}^*)$ . Then, for all  $D > \underline{D}$ , we have that  $\pi_i(G, x_i^*, \mathbf{x}_{-i}^*) > -D$ , namely, defeat is not a profitable deviation. This completes the proof of existence of an interior Nash Equilibrium.

Next, we prove uniqueness. We assume that, contrary to the statement of the proposition, for all  $D < \infty$ , there exists an equilibrium where  $n - \hat{n} > 0$  groups take the defeat option. Then, we show that this induces a contradiction. Since we have proved that when all  $n$  groups participate in the contest there exists a unique equilibrium, this establishes global uniqueness.

The condition that  $\Gamma_i^{0, \gamma} > 0, \forall i = 1, 2, \dots, n$ , ensures that, in a candidate equilibrium in which only  $\hat{n} < n$  groups participate in the contest, all such  $\hat{n}$  groups choose a finite effort level (this follows immediately from the analysis of the case where all  $n$  groups participate). The effort level of participants is  $\mathbf{x}_{\hat{n}}^* = \Lambda^{\beta, \gamma}(G_{\hat{n}})(1 - \Lambda^{\beta, \gamma}(G_{\hat{n}}))\mathbf{c}^{\beta, \gamma}(G_{\hat{n}})$  where the graph  $G_{\hat{n}}$  only includes the participating groups. Consider a non-participating group  $\nu$ . For this group, in the assumed equilibrium,  $\pi_\nu = -D$ . Suppose group  $\nu$  deviates and chooses, instead,  $x_\nu = x_\nu^0$ , where  $x_\nu^0$  is the unique threshold such that  $\varphi_\nu(G_{\hat{n}+1}, x_\nu^0, \hat{\mathbf{x}}_{-\nu}^*) = 0$ . The payoff of this deviation is  $\pi_\nu(G_{\hat{n}+1}, x_\nu^0, \hat{\mathbf{x}}_{-\nu}^*) = -x_\nu^0 > -\infty$ . Thus, for any  $D > x_\nu^0$ , this deviation is profitable. Repeating the argument for all partitions establishes that there exists  $\underline{D} < \infty$  such that, for all  $D > \underline{D}$ , any candidate equilibrium where  $n - \hat{n} > 0$  groups take the defeat option is susceptible to a profitable deviation (hence, it is not an equilibrium). Thus, the only equilibrium is interior, completing the proof.

**Remark:** Figure A.1 (referred to in the text) shows the payoff function  $\pi_i(G, [x_i, \mathbf{x}_{-i}^*])$  at the equilibrium strategy profile. Group  $i$ 's payoff function is constant ( $\pi_i = -D$ ) for all  $x_i$  below the threshold that guarantees the non-negativity of  $\varphi_i$ . At the threshold, the function is discontinuous, capturing the fact that when  $\varphi_i \geq 0$  no defeat cost is due. To the right of the threshold, condition (3) ensures that  $\pi_i(G, [x_i, \mathbf{x}_{-i}^*])$  is strictly concave in  $x_i$ . Moreover, the payoff function is hump shaped and reaches a maximum at  $\varphi_i^* > 0$ .

## A.2 THE CASE OF SMALL EXTERNALITIES

The centrality measure  $c_i^{\beta,\gamma}(G)$  in equation (8) depends, in general, on the entire network structure. However, it is instructive to consider networks in which the spillover parameters  $\beta$  and  $\gamma$  are small. In this case, our centrality measure can be approximated by the sum of (i) the Katz-Bonacich centrality related to the network of enmities,  $G^-$ , (ii) the (negative-parameter) Katz-Bonacich centrality related to the network of alliances,  $G^+$ , and (iii) the local hostility vector,  $\mathbf{\Gamma}^{\beta,\gamma}(G)$ .<sup>39</sup>

**Lemma 1.** *Assume that the conditions of Proposition 1 are satisfied. Then, as  $\beta \rightarrow 0$  and  $\gamma \rightarrow 0$ , the centrality measure defined in equation (8) can be written as*

$$\mathbf{c}^{\beta,\gamma}(G) = \mathbf{b}(\gamma, G^-) + \mathbf{b}(-\beta, G^+) - \mathbf{\Gamma}^{\beta,\gamma}(G) + O(\beta\gamma),$$

where  $O(\beta\gamma)$  involves second and higher order terms, and the (weighted) Katz-Bonacich centrality with parameter  $\alpha$  is defined as  $\mathbf{b}(\alpha, G) \equiv \mathbf{b}_{\mathbf{\Gamma}^{\beta,\gamma}(G)}(\alpha, G) = (\mathbf{I}_n - \alpha \mathbf{A})^{-1} \mathbf{\Gamma}^{\beta,\gamma}(G)$  assuming that  $\alpha$  is smaller than the inverse of the largest eigenvalue of  $\mathbf{A}$ .

Lemma 3 states that the centrality  $\mathbf{c}^{\beta,\gamma}(G)$  can be expressed as a linear combination of the weighted Katz-Bonacich centralities  $\mathbf{b}(\gamma, G^-)$ ,  $\mathbf{b}(-\beta, G^+)$  and the vector  $\mathbf{\Gamma}^{\beta,\gamma}(G)$ . Each Katz-Bonacich centrality gauges the network multiplier effect attached to the system of enmities and alliances, respectively. In the case of weak network externalities (i.e. when  $\beta \rightarrow 0$  and  $\gamma \rightarrow 0$ ), the following approximation holds true:

$$b_{\cdot,i}(\gamma, G^-) = \Gamma_i^{\beta,\gamma}(G) + \gamma \sum_{j=1}^n a_{ij}^- \Gamma_j^{\beta,\gamma}(G) + \gamma^2 \sum_{j=1}^n a_{ij}^- \sum_{k=1}^n a_{jk}^- \Gamma_k^{\beta,\gamma}(G) + O(\gamma^3), \quad (28)$$

$$b_{\cdot,i}(-\beta, G^+) = \Gamma_i^{\beta,\gamma}(G) + (-\beta) \sum_{j=1}^n a_{ij}^+ \Gamma_j^{\beta,\gamma}(G) + (-\beta)^2 \sum_{j=1}^n a_{ij}^+ \sum_{k=1}^n a_{jk}^+ \Gamma_k^{\beta,\gamma}(G) + O(\beta^3). \quad (29)$$

Thus, Lemma 3 suggests that, when higher order terms can be neglected, our centrality measure is increasing in  $\gamma$  and in the number of first- and second-degree enmities, whereas it is decreasing in  $\beta$  and in the number of first-degree alliances. Second-degree alliances have instead a positive effect on the centrality measure. An illustration of this result in the case of a path graph is provided in Appendix C.

We can also obtain a simple approximate expression for the equilibrium efforts and the payoffs in Proposition 1.<sup>40</sup>

**Lemma 2.** *As  $\beta \rightarrow 0$  and  $\gamma \rightarrow 0$ , the equilibrium effort and payoff of group  $i$  in network  $G$  can be written as*

$$\begin{aligned} x_i^*(G) &= A_1^{\beta,\gamma}(G) - B_1 (\beta d_i^+ - \gamma d_i^-) + O(\beta\gamma), \\ \pi_i^*(G) &= A_2^{\beta,\gamma}(G) + B_2 (\beta d_i^+ - \gamma d_i^-) + O(\beta\gamma), \end{aligned}$$

where  $A_1^{\beta,\gamma}(G)$ ,  $B_1$ ,  $A_2^{\beta,\gamma}(G)$  and  $B_2$  are positive constants with  $A_1^{\beta,\gamma}(G)$  and  $A_2^{\beta,\gamma}(G)$  being of the order of  $O(\beta) + O(\gamma)$ .

Lemma 4 shows that, when network externalities are small, a group's fighting effort increases in the weighted difference between the number of enmities (weighted by  $\gamma$ ) and alliances (weighted by  $\beta$ ), i.e. the net local externalities  $d_i^- \gamma - d_i^+ \beta$ . The opposite is true for the equilibrium payoff, which is increasing in  $d_i^+ \beta - d_i^- \gamma$ . Thus, ceteris paribus, an increase in the spillover from alliances (enmities), parameterized by  $\beta$  ( $\gamma$ ), and an increase in the number of allies (enemies) decreases (increases) group  $i$ 's fighting effort and

<sup>39</sup>See Appendix C for a formal definition of the Katz-Bonacich centrality and for proofs of Lemmas 1-2.

<sup>40</sup>See the proof of Lemma 4 for the explicit expressions for the constants  $A^{\beta,\gamma}(G)$ ,  $B$ ,  $C^{\beta,\gamma}(G)$  and  $D$ .

It is useful to note that, when  $\beta = \gamma = 0$ , then  $\Lambda^{\beta,\gamma}(G) = 1 - \frac{1}{n}$ , and  $c_i^{\beta,\gamma}(G) = 1$ . Then, the equilibrium expressions in Proposition 1 simplify to  $x_i^*(G) = (n-1)/n^2$  and  $\pi_i^*(G) = 1/n^2$  which are the standard solutions in the Tullock CSF.

increases (reduces) its payoff. Intuitively, a group with many enemies tends to fight harder and to appropriate a smaller share of the prize, whereas a group with many friends tends to fight less and to appropriate a large size of the prize. One must remember, however, that this simple result hinges on  $\beta$  and  $\gamma$  being small; in general, higher-degree links have sizeable effects.

### A.3 APPENDIX TO SECTION 2.4 (HETEROGENEITY)

In the following we provide a complete equilibrium characterization of the extension of our model that we have introduced in Section 2.4. When the fighting strength  $\varphi_i$  of group  $i$  depends on an idiosyncratic shifter  $\tilde{\varphi}_i$  as in equation (11) then the following proposition characterizes the corresponding Nash equilibrium.

**Proposition 2.** *Let  $\Gamma_i^{\beta,\gamma}(G)$  and  $\Lambda^{\beta,\gamma}(G)$  be defined as in equation (6), and let*

$$\mathbf{c}_{\boldsymbol{\mu}}^{\beta,\gamma}(G) \equiv (\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-)^{-1} \boldsymbol{\mu} \quad (30)$$

be a centrality vector, whose generic element  $c_{\boldsymbol{\mu},i}^{\beta,\gamma}(G)$  describes the centrality of group  $i$  in the network for some vector  $\boldsymbol{\mu} \in \mathbb{R}^n$ , and assume that the same parameter restrictions on  $\beta$  and  $\gamma$  hold as in Proposition 1. Then for the cost of defeat,  $D$ , large enough there exists a unique Nash equilibrium of the  $n$ -player simultaneous move game with payoffs given by equation (1), groups' OPs in equation (11) and strategy space  $\mathbb{R}^n$ , where the equilibrium effort levels are given by

$$x_i^*(G) = \Lambda^{\beta,\gamma}(G) (1 - \Lambda^{\beta,\gamma}(G)) c_{\Gamma^{\beta,\gamma}(G),i}^{\beta,\gamma}(G) - c_{\tilde{\varphi},i}^{\beta,\gamma}(G), \quad (31)$$

for all  $i = 1, \dots, n$ . Moreover, the aggregate and individual equilibrium OPs are, respectively,

$$\sum_{i=1}^n \varphi_i = \Lambda^{\beta,\gamma}(G), \quad (32)$$

$$\varphi_i = \Lambda^{\beta,\gamma}(G) (1 - \Lambda^{\beta,\gamma}(G)) \Gamma_i^{\beta,\gamma}(G), \quad (33)$$

and the equilibrium payoffs are given by

$$\pi_i^*(G) \equiv \pi_i(\mathbf{x}^*, G) = (1 - \Lambda^{\beta,\gamma}(G)) \left( \Gamma_i^{\beta,\gamma}(G) - \Lambda^{\beta,\gamma}(G) c_{\Gamma^{\beta,\gamma}(G),i}^{\beta,\gamma}(G) \right) + c_{\tilde{\varphi},i}^{\beta,\gamma}(G). \quad (34)$$

**Proof of Proposition 2.** We start by considering an equilibrium in which all groups participate in the contest (i.e. an interior equilibrium). With equation (11) we can write group  $i$ 's payoff as follows

$$\begin{aligned} \pi_i(G, \mathbf{x}) &= \frac{\varphi_i}{\sum_{j=1}^n \varphi_j} - x_i \\ &= \frac{x_i + \beta \sum_{j=1}^n a_{ij}^+ x_j - \gamma \sum_{j=1}^n a_{ij}^- x_j + \tilde{\varphi}_i}{\sum_{j=1}^n \left( x_j + \beta \sum_{k=1}^n a_{jk}^+ x_k - \gamma \sum_{k=1}^n a_{jk}^- x_k + \tilde{\varphi}_j \right)} - x_i. \end{aligned} \quad (35)$$

The partial derivatives are given by

$$\begin{aligned} \frac{\partial \pi_i}{\partial x_i} &= \frac{\frac{\partial \varphi_i}{\partial x_i} \sum_{j=1}^n \varphi_j - \varphi_i \sum_{j=1}^n \frac{\partial \varphi_j}{\partial x_i}}{\left( \sum_{j=1}^n \varphi_j \right)^2} - 1 \\ &= \frac{\sum_{j=1}^n \varphi_j - \varphi_i (1 + \beta d_i^+ - \gamma d_i^-)}{\left( \sum_{j=1}^n \varphi_j \right)^2} - 1, \end{aligned} \quad (36)$$

where we have used the fact that  $\frac{\partial \varphi_i}{\partial x_i} = \delta_{ij} + \beta a_{ij}^+ - \gamma a_{ij}^-$  and consequently  $\sum_{j=1}^n \frac{\partial \varphi_j}{\partial x_i} = 1 + \beta d_i^+ - \gamma d_i^-$ . The FOCs are then given by  $\frac{\partial \pi_i}{\partial x_i} = 0$ . From the partial derivative in equation (36) the FOC for group  $i$  can be written as follows

$$\frac{\partial \pi_i}{\partial x_i} = \frac{\sum_{j=1}^n \varphi_j - \varphi_i(1 + \beta d_i^+ - \gamma d_i^-)}{\left(\sum_{j=1}^n \varphi_j\right)^2} - 1 = 0,$$

from which we get

$$\varphi_i = \frac{1}{1 + \beta d_i^+ - \gamma d_i^-} \left(1 - \sum_{j=1}^n \varphi_j\right) \sum_{j=1}^n \varphi_j.$$

Summation over  $i$  gives

$$\sum_{i=1}^n \varphi_i = \left(1 - \frac{1}{\sum_{i=1}^n \frac{1}{1 + \beta d_i^+ - \gamma d_i^-}}\right).$$

With  $\Gamma_i^{\beta, \gamma}(G)$  and  $\Lambda^{\beta, \gamma}(G)$  as in equation (6) we can write the aggregate operational performance as

$$\sum_{i=1}^n \varphi_i = \Lambda^{\beta, \gamma}(G),$$

which is equivalent to equation (32). The individual operational performance can be written as

$$\varphi_i(G, \mathbf{x}) = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G))\Gamma_i^{\beta, \gamma}(G), \quad (37)$$

which is equivalent to equation (32). We then get

$$\varphi_i(G, \mathbf{x}) = x_i + \beta \sum_{j=1}^n a_{ij}^+ x_j - \gamma \sum_{j=1}^n a_{ij}^- x_j + \tilde{\varphi}_i = \Lambda^{n, \beta}(G)(1 - \Lambda^{n, \beta}(G))\Gamma_i^{\beta, \gamma}(G). \quad (38)$$

We can write

$$x_i + \beta \sum_{j=1}^n a_{ij}^+ x_j - \gamma \sum_{j=1}^n a_{ij}^- x_j = \varphi_i - \tilde{\varphi}_i = \Lambda^{n, \beta}(G)(1 - \Lambda^{n, \beta}(G))\Gamma_i^{\beta, \gamma}(G) - \tilde{\varphi}_i.$$

Denoting by  $\Gamma^{\beta, \gamma}(G) \equiv (\Gamma_1^{\beta, \gamma}(G), \dots, \Gamma_n^{\beta, \gamma}(G))^T$ , we can write this in vector-matrix form as

$$(\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-) \mathbf{x} = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G))\mathbf{\Gamma}^{\beta, \gamma}(G) - \tilde{\boldsymbol{\varphi}}.$$

When the matrix  $\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-$  is invertible, we obtain a unique solution given by

$$\mathbf{x} = \Lambda^{\beta, \gamma}(G)(1 - \Lambda^{\beta, \gamma}(G))(\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-)^{-1} \mathbf{\Gamma}^{\beta, \gamma}(G) - (\mathbf{I}_n + \beta \mathbf{A}^+ - \gamma \mathbf{A}^-)^{-1} \tilde{\boldsymbol{\varphi}}. \quad (39)$$

With the definition of the centrality in equation (30) we then can write equation (39) in the form of equation (31) in the proposition. Moreover, using the fact that equilibrium payoffs are given by  $\pi_i(x^*, G) = \frac{\varphi_i^*(G)}{\sum_{j=1}^n \varphi_j^*(G)} - x_i^*$  we obtain equation (34) in the proposition.

Next, note that from equation (36) we find that

$$\frac{\partial^2 \pi_i}{\partial x_i^2} = -\frac{2}{\Lambda^{\beta, \gamma}(G)\Gamma_i^{\beta, \gamma}(G)} = -2 \frac{\sum_{j=1}^n \Gamma_j^{\beta, \gamma}(G)}{\sum_{j=1}^n \Gamma_j^{\beta, \gamma}(G) - 1} (1 + \beta d_i^+ - \gamma d_i^-), \quad (40)$$

which is negative under the same parameter restrictions on  $\beta$  and  $\gamma$  as in Proposition 1. Further, as in the proof of Proposition 1 one can then proceed to establish that there exists a unique interior Nash equilibrium when the cost of defeat,  $D$ , is large enough. ■

## B EMPIRICAL APPENDIX

In this section, we discuss technical details related to Sections 4 (ambiguous links) and 6 (random utility model). Then, we include the tables and figures discussed in the main text and in this appendix. Additional tables and figures can be found in the webpage Appendix C.

### B.1 APPENDIX TO SECTION 4.2.2

While the large majority of bilateral links are incontrovertible, there are some that may potentially be ambiguous, i.e. for which experts do not all reach the same clear-cut assessments or where the battleground observations in ACLED are not fully reflected in the expert accounts. As mentioned earlier, Table B.5 checks the robustness of our results with respect to ambiguous group links. The columns 1–5 of the table have already been discussed in the main text. Here, we discuss columns 6–9. In column 6, we exclude the CNDD since this group appears to have switched its relation with the Mayi-Mayi militia. Next, we classify Uganda and the RCD-G as enemies (column 7). While this violates our coding rule (that classifies them as neutral), it is more consistent with the narrative. Next, we code all member states of the Southern African Development Community (SADC) as allies of each other and of the FARDC (column 8). Finally, we define as “governments allied to the FARDC” all governments allied to the FARDC in the baseline treatment plus all SADC member states. Finally, we let all “governments allied to the FARDC” be (i) allied among themselves, (ii) allied to the FARDC, and (iii) enemies to Rwanda-I and Uganda (column 9). The results are in all cases similar to the baseline table.

In Table B.6 we investigate potentially ambiguous links on the basis of accounts in Auteserre (2008), Prunier (2011), Sanchez de la Sierra (2017), Stearns (2011). We start by focusing in column 1 on the FARDC links with ALIR. In the baseline regression, they are coded as enemies under Laurent Kabila and as neutral under Joseph Kabila. The earlier link is potentially ambiguous, since the ALIR was part of the Hutu exodus that clashed with Rwanda and Uganda (see, e.g., Prunier 2011: 234). This might suggest an alliance between the FARDC and ALIR. However, this is in contradiction with the ACLED data where these two groups are observed fighting against each other on fourteen occasions between 1998 and 2001, and never on the same side as brothers in arms. Column 1 shows that the results are not sensitive if we code this dyad as an alliance or as an enmity.

In the same vein, columns 2 and 3 run the same regressions as in columns 4 and 5 of Table B.5 where we investigated the ambiguous links between the FARDC and FDLR, but assuming now in addition the same links for the FARDC-Interahamwe pair as for the FARDC-FDLR pair. In particular, in our baseline estimates the FARDC-FDLR are first allies under Laurent Kabila, and then neutral under Joseph Kabila, whereas the FARDC-Interahamwe are allies throughout the whole period. Here, we assume that FARDC-Interahamwe and FARDC-FDLR are in both cases allies under Laurent Kabila and enemies under Joseph Kabila (column 2), or neutral throughout the entire period (column 3).

In column 4 we scrutinize the relationship between the FARDC-LK and ADF. In the baseline we code them as allied, based on an expert source (see International Crisis Group, 1998). However, Prunier (2011: 177) mentions one incident where the 10th division of the FARDC attacks ADF rebels. While this appears to be an isolated episode, in column 4 we code FARDC-LK and ADF as enemies.

In column 5, we investigate the link MLC-Hema ethnic militia. In the baseline estimations, we code them as enemy, based on ACLED reporting them to fight five times against each other and never as brothers in arms. Stearns (2011: 230) refers to one incident where MLC troops cooperated with the Hema ethnic militia to attack locals. Thus, in the robustness check of column 5 we code them as allies.

Finally, the link FDLR-MayiMayi is considered in column 6. In the baseline estimations we code two MayiMayi fractions as allied to FDLR (based on ACLED incidences), namely the Yakutumba and Pareco fractions. In contrast, the relationship FDLR-MayiMayi is coded as neutral for all other MayiMayi fractions (again based on ACLED). Prompted by Sanchez de la Sierra’s statement that “Mayi-Mayi groups and the FDLR often conducted joint operations during the second Congo war” (2016: 84), we code all MayiMayi fractions as being allied to the FDLR.

In all columns the baseline results are virtually unchanged. The results are also robust to changing all links simultaneously (except for column 3, since it would be incompatible with column 2).

## B.2 APPENDIX TO SECTION 6 (RANDOM UTILITY MODEL)

### B.2.1 Multinomial logit estimation

Table B.12 displays the results of the multinomial logit estimation. Table B.13 displays the results of the Random Utility Model with time-varying network. Figure B.1 reports the cross-dyad distribution of predicted probabilities for enmities (left panel) and alliances (right panel). In the left panel, the dark sample represents the distribution for observed enmities (alliances) and the light sample represents the distribution for the other observed links. Figure B.2 displays the observed and predicted distributions of six network statistics that play an important role in our baseline model: degree one enemies (panel *a*), degree one allies (panel *b*), number of degree one links (panel *c*), common enemies (panel *d*), common allies (panel *e*), and conflicting neighbors (panel *f*). For each statistic, the panel compares the data with the average distribution over 1000 simulated networks as predicted by the conditional logit model (plus/minus one standard deviation interval). The figure shows that all these important moments of the data are predicted accurately. More precisely each Monte Carlo draw consists in drawing a random utility shock  $\tilde{u}_{ij}(a)$  for each dyad  $ij$  and alternative  $a$  from a type I extreme value distribution with mean  $\gamma \approx 0.577$  (the Euler-Mascheroni constant), and variance,  $\sqrt{\pi}/6$ . Using equation (21) we compute each simulated alternative-dependent joint surplus  $U_{ij}^{sim}(a)$  and select the predicted link for dyad  $ij$  as  $a_{ij}^{pre} = \arg \max U_{ij}^{sim}(a)$ . This leads to a simulated network for which we can compute the 6 distributions of network statistics. This procedure is iterated 1000 times. Each panel reports the average distribution ( $\pm 1$  SD) across the Monte Carlo draws.

### B.2.2 Effect of removing armed groups (with network rewiring)

Table B.14 summarizes the effect of removing one or more groups from the conflict when endogenous network adjustments are allowed. We focus, for comparability, on the top 15 groups in Table 4. All but three groups remain in the top 15 even after allowing for network recomposition.<sup>41</sup> The UPC is especially interesting. This is a medium-large group whose activity accounts for 2.2% of the total violence. Its removal in Table 4 yields a reduction in violence of the order of 3%, with a sizable multiplier of 1.4. However, the recomposition of the network after its removal offsets two thirds of the gains.

Table B.1: Summary statistics.

Variable	Obs.	Mean	Std. Dev.	Min	Max
Total Fight.	1,040	5.929	25.046	0	300
Total Fight. Enemies (TFE)	1,040	69.237	109.95	0	682
Total Fight. Allies (TFA)	1,040	48.603	85.75	0	563
Total Fight. Neutrals (TFN)	1,040	350.539	241.616	1	1042
$d^-$ (# enemies)	1,040	2.95	4.306	0	26
$d^+$ (# allies)	1,040	2.4	3.45	0	21
Rainfall ( $t - 1$ )	1,040	125.839	26.164	59.639	195.56

Note: The sample comprises the 80 fighting groups that are involved in at least one fighting event in ACLED during the period 1998–2010.

<sup>41</sup>The three groups that drop out of the top 15 are the UPC, Mutiny of FARC, and FRPI. The group entering the top 15 are two branches of the RCD (the first collects all events involving "unspecified" RCD; the second is the group labeled RCD-National; both are likely to suffer with large measurement error) and the National Army for the Liberation of Uganda.

Table B.2: Benchmark second stage with unpopular<sub>6</sub>.

	Dependent variable: Total fighting							
	OLS	Reduced IV	Full IV	Neutrals	Battles	$d^- \geq 1$ & $d^+ \geq 1$	GED coord.	GED union
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Fight. Enemies (TFE)	0.066*** (0.016)	0.130** (0.057)	0.100*** (0.033)	0.119*** (0.035)	0.121*** (0.035)	0.118*** (0.035)	0.123*** (0.036)	0.161*** (0.045)
Total Fight. Allies (TFA)	0.001 (0.017)	-0.218** (0.086)	-0.149*** (0.051)	-0.139*** (0.047)	-0.153*** (0.051)	-0.161*** (0.062)	-0.138*** (0.047)	-0.152*** (0.050)
Total Fight. Neutrals (TFN)				0.005 (0.006)	0.004 (0.006)	0.009 (0.014)	0.005 (0.006)	0.006 (0.005)
Additional controls	Reduced	Reduced	Full	Full	Full	Full	Full	Full
Estimator	OLS	IV	IV	IV	IV	IV	IV	IV
Set of Instrument Variables	n.a.	Restricted	Full	Full	Full	Full	Full	Full
Kleibergen-Paap F-stat	n.a.	10.6	15.6	15.1	17.4	14.9	15.1	17.3
Hansen J (p-value)	n.a.	0.16	0.75	0.83	0.79	0.70	0.81	0.79
Observations	1040	1040	1040	1040	988	598	1040	1781
R-squared	0.510	0.265	0.474	0.459	0.446	0.462	0.455	0.389

Note: All regressions include group fixed effects and control for rainfall in the own group's territory. Columns 1-3 include time fixed effects. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.3: Robustness.

	Dependent variable: Total fighting									
	Only deg. 2 (1)	Sample split (2)	Only violent (3)	Batt., riots, viol. (4)	Excl. bilateral (5)	Lags (6)	1 evt. enemy (7)	2 evts. ally (8)	RUM (9)	3d Poly. (10)
Total Fight. Enemies (TFE)	0.074*** (0.027)	0.147** (0.062)	0.083*** (0.019)	0.083*** (0.019)	0.063*** (0.023)	0.090*** (0.026)	0.119*** (0.029)	0.089*** (0.019)	0.106*** (0.033)	0.082*** (0.019)
Total Fight. Allies (TFA)	-0.207** (0.083)	0.002 (0.079)	-0.114*** (0.033)	-0.114*** (0.034)	-0.110*** (0.033)	-0.104*** (0.040)	-0.170*** (0.051)	-0.107*** (0.033)	-0.112** (0.050)	-0.115*** (0.034)
Total Fight. Neutrals (TFN)	0.008 (0.006)	0.011 (0.007)	0.004 (0.004)	0.004 (0.004)	0.004 (0.004)	0.007 (0.005)	0.005 (0.006)	0.002 (0.004)	0.002 (0.006)	0.004 (0.004)
Additional controls	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full
Estimator	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
Set of Instrument Variables	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full
Kleibergen-Paap F-stat	13.4	37.7	22.5	24.6	20.4	20.3	11.8	25.3	4.0	28.6
Hansen J (p-value)	0.28	0.84	0.58	0.58	0.54	0.69	0.51	0.51	0.62	0.58
Observations	1040	640	1040	1027	1040	960	1040	1040	1040	1040
R-squared	0.489	0.589	0.568	0.570	0.550	0.577	0.394	0.571	0.554	0.567

Note: All regressions include group fixed effects and control for rainfall in the own group's territory. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



Table B.4: Robustness: Group definition (FARDC, Rwanda & others).

Dependent variable: Total fighting				
Merge	FARDC	MayMay	Rwanda	FARDC & RWA
	(1)	(2)	(3)	(4)
Total Fight. Enemies (TFE)	0.040** (0.019)	0.068*** (0.022)	0.074*** (0.022)	0.040** (0.020)
Total Fight Allies (TFA)	-0.058* (0.034)	-0.084*** (0.032)	-0.108*** (0.032)	-0.066* (0.037)
Total Fight. Neutrals (TFN)	-0.002 (0.004)	-0.000 (0.005)	0.003 (0.005)	-0.003 (0.004)
Additional controls	Full	Full	Full	Full
Estimator	IV	IV	IV	IV
Set of Instrument Variables	Full	Full	Full	Full
Kleibergen-Paap F-stat	34.6	15.4	19.7	29.8
Hansen J (p-value)	0.44	0.66	0.59	0.51
Observations	1027	962	1027	1014
R-squared	0.631	0.614	0.589	0.642

Note: All regressions include group fixed effects and control for rainfall in the group's homeland. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.5: Robustness: Ambiguous network links I.

Dependent variable: Total fighting									
	Uga-Rwa neutral (1)	Uga-Rwa allies (2)	Uga-Rwa allies then neutral (3)	FARDC-FDLR allies then enemies (4)	FARDC-FDLR neutral (5)	Exclude CNDD (6)	Uga-RCD-G enemies (7)	SADC allies (8)	Dyadic closure main groups (9)
Total Fight. Enemies (TFE)	0.082*** (0.021)	0.081*** (0.021)	0.082*** (0.020)	0.090*** (0.020)	0.083*** (0.020)	0.083*** (0.020)	0.074*** (0.019)	0.082*** (0.020)	0.075*** (0.020)
Total Fight Allies (TFA)	-0.105*** (0.031)	-0.109*** (0.033)	-0.112*** (0.033)	-0.116*** (0.033)	-0.112*** (0.036)	-0.111*** (0.033)	-0.104*** (0.031)	-0.116*** (0.035)	-0.100*** (0.031)
Total Fight. Neutrals (TFN)	0.004 (0.004)	0.003 (0.004)	0.004 (0.004)	0.003 (0.004)	0.003 (0.004)	0.002 (0.004)	0.004 (0.004)	0.004 (0.004)	0.000 (0.004)
Additional controls	Full	Full	Full	Full	Full	Full	Full	Full	Full
Estimator	IV	IV	IV	IV	IV	IV	IV	IV	IV
Set of Instrument Variables	Full	Full	Full	Full	Full	Full	Full	Full	Full
Kleibergen-Paap F-stat	19.9	20.2	19.8	16.7	23.4	27.2	20.8	19.2	12.7
Hansen J (p-value)	0.60	0.59	0.58	0.60	0.56	0.54	0.57	0.55	0.58
Observations	1040	1040	1040	1040	1040	1027	1040	1040	1040
R-squared	0.574	0.573	0.569	0.577	0.565	0.574	0.577	0.573	0.589

Note: All regressions include group fixed effects and control for rainfall in the group's homeland. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.6: Robustness: Ambiguous network links II.

	Dependent variable: Total fighting					
	FARDC-ALIR ally (1)	FARDC-FDLR/Interah. ally then enemy (2)	FARDC-FDLR/Interah. neutral (3)	FARDC-ADF enemy (4)	Hema ethnic militia-MLC ally (5)	FDLR-MayiMayi ally (6)
Total Fight. Enemies (TFE)	0.087*** (0.020)	0.077*** (0.022)	0.078*** (0.019)	0.081*** (0.018)	0.085*** (0.019)	0.070*** (0.020)
Total Fight. Allies (TFA)	-0.117*** (0.035)	-0.087*** (0.025)	-0.103*** (0.034)	-0.117*** (0.036)	-0.113*** (0.035)	-0.107*** (0.033)
Total Fight. Neutrals (TFN)	0.004 (0.004)	0.002 (0.004)	0.003 (0.004)	0.004 (0.004)	0.004 (0.004)	0.003 (0.006)
Additional controls	Full	Full	Full	Full	Full	Full
Estimator	IV	IV	IV	IV	IV	IV
Set of Instrument Variables	Full	Full	Full	Full	Full	Full
Observations	1040	1040	1040	1040	1040	1040
Kleibergen-Paap F-stat	21.4	10.1	25.2	25.1	20.1	28.3
Hansen J (p-value)	0.58	0.58	0.55	0.56	0.57	0.58
R-squared	0.562	0.605	0.576	0.570	0.566	0.574

Note: All regressions include group fixed effects and control for rainfall in the own group's territory. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.7: Measurement error in rainfall.

Dependent variable: GPCC gauge rainfall measure								
Model	Linear		Log-linear		Linear		Log-linear	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# ACLED conflict events	0.017 (0.032)	0.008 (0.012)	0.009 (0.008)	0.001 (0.003)	-0.069 (0.057)	0.016 (0.014)	-0.016 (0.014)	0.005 (0.004)
TRMM satellite rainfall measure	0.639*** (0.018)	0.513*** (0.012)	0.714*** (0.015)	0.619*** (0.013)	- -	- -	- -	- -
GPCP satellite rainfall measure	- -	- -	- -	- -	0.790*** (0.044)	1.073*** (0.081)	0.843*** (0.055)	1.233*** (0.100)
(0.5 x 0.5) Grid Cell FE	No	Yes	No	Yes	No	Yes	No	Yes
Annual TE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9893	9893	9893	9893	9893	9893	9893	9893
R-squared	0.578	0.667	0.604	0.684	0.555	0.601	0.541	0.587

Note: The unit of observation is a cell of resolution 0.5 x 0.5 degrees in a given year. The panel contains 761 cells covering DRC between 1998 and 2010. In Columns 3,4,5,6 all rainfall variables are in log. Robust standard errors are clustered at the (0.5 x 0.5) cell level in Columns 1-4 and at the (2.5 x 2.5) cell level in Columns 5-8. Significance levels are indicated by \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01.

Table B.8: Alternative rainfall data.

Panel a (TRMM)						
Column in baseline Table 1	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)
Definition “large group”	$d^- \geq 10$		$d^- \geq 6$		not included	
	(1)	(2)	(3)	(4)	(5)	(6)
Total Fight. Enemies (TFE)	0.025 (0.029)	0.084** (0.042)	0.053* (0.029)	0.102** (0.049)	0.131** (0.060)	0.197*** (0.074)
Total Fight. Allies (TFA)	-0.092** (0.038)	-0.146*** (0.047)	-0.142*** (0.053)	-0.194*** (0.058)	-0.126** (0.058)	-0.187*** (0.066)
Total Fight. Neutrals (TFN)	0.006 (0.005)	0.007 (0.008)	0.005 (0.007)	0.007 (0.009)	0.004 (0.006)	0.003 (0.010)
Kleibergen-Paap F-stat	10.3	3.9	10.9	4.2	18.1	10.7
Hansen J (p-value)	0.50	0.68	0.55	0.72	0.39	0.72
Observations	1040	988	1040	988	1040	988
R-squared	0.586	0.534	0.495	0.403	0.431	0.264
Panel b (GPCP)						
Column in baseline Table 1	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)
Definition “large group”	$d^- \geq 10$		$d^- \geq 6$		not included	
	(1)	(2)	(3)	(4)	(5)	(6)
Total Fight. Enemies (TFE)	0.017 (0.041)	0.023 (0.041)	0.060 (0.050)	0.064 (0.053)	0.138** (0.063)	0.155** (0.062)
Total Fight. Allies (TFA)	-0.068*** (0.026)	-0.097*** (0.036)	-0.130*** (0.049)	-0.186*** (0.062)	-0.106* (0.055)	-0.144** (0.064)
Total Fight. Neutrals (TFN)	-0.009 (0.008)	-0.006 (0.009)	-0.010 (0.011)	-0.004 (0.013)	-0.002 (0.012)	0.001 (0.013)
Kleibergen-Paap F-stat	7.2	7.4	6.4	11.3	10.1	10.0
Hansen J (p-value)	0.60	0.69	0.66	0.72	0.58	0.71
Observations	1040	988	1040	988	1040	988
R-squared	0.606	0.590	0.514	0.446	0.444	0.390
Panel c (TRMM & GPCP)						
Column in baseline Table 1	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)	4 (all evts.)	5 (battles)
Definition “large group”	$d^- \geq 10$		$d^- \geq 6$		not included	
	(1)	(2)	(3)	(4)	(5)	(6)
Total Fight. Enemies (TFE)	0.035 (0.026)	0.060** (0.026)	0.069** (0.029)	0.095*** (0.033)	0.128*** (0.049)	0.163*** (0.050)
Total Fight Allies (TFA)	-0.063*** (0.024)	-0.099*** (0.028)	-0.108*** (0.038)	-0.157*** (0.043)	-0.087* (0.046)	-0.129*** (0.050)
Total Fight. Neutrals (TFN)	0.006 (0.005)	0.007 (0.007)	0.002 (0.006)	0.004 (0.008)	0.003 (0.006)	0.001 (0.009)
Kleibergen-Paap F-stat	38.6	27.8	51.4	33.3	58.0	15.6
Hansen J (p-value)	0.89	0.93	0.93	0.94	0.85	0.87
Observations	1040	988	1040	988	1040	988
R-squared	0.609	0.586	0.528	0.464	0.471	0.398

Note: The dependent variable is total fighting. The unit of observation is an armed group in a given year. The panel contains 80 armed groups between 1998 and 2010. All regressions include group fixed effects and control for rainfall in the own group’s territory, as well as the same additional controls and instruments as in the columns (4) and (5) of the baseline Table 1. Robust standard errors corrected for Spatial HAC in parentheses. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.9: Monte Carlo simulations testing link mismeasurement.

Probability of mismeasurement		0		0.01		0.1		0.2		0.5		1	
		TFA	TFE	TFA	TFE	TFA	TFE	TFA	TFE	TFA	TFE	TFA	TFE
Enmity links only	Mean	0.118	0.085	0.117	0.084	0.112	0.077	0.106	0.0713	0.102	0.046	0.119	0.002
	S.D.	0.012	0.011	0.012	0.011	0.014	0.013	0.019	0.022	0.024	0.020	0.018	0.015
Alliance links only	Mean	0.119	0.085	0.117	0.0843	0.107	0.075	0.095	0.066	0.059	0.055	0.000	0.083
	S.D.	0.012	0.011	0.014	0.011	0.019	0.015	0.023	0.019	0.027	0.022	0.008	0.008
Alliance & Enmity links	Mean	0.118	0.085	0.116	0.083	0.101	0.067	0.086	0.051	0.040	0.012	0.001	-0.001
	S.D.	0.012	0.010	0.014	0.012	0.022	0.019	0.028	0.023	0.037	0.028	0.017	0.014

Note: The table reports the mean and standard deviations (S.D.) of the Monte Carlo sampling distributions (1,000 draws) of the baseline 2SLS estimates (col. 4, Table 1) of TFE and TFA for different probabilities of network link mismeasurement. The data generating process is based on the coefficients = 0.114 for TFA and = 0.083 for TFE.

Table B.10: Bootstrapping ACLED Events

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFA	Mean	0.006	-0.231	-0.121	-0.114	-0.113	-0.138	-0.111	-0.132
	S.D.	0.007	0.039	0.022	0.023	0.023	0.033	0.023	0.013
TFE	Mean	0.065	0.131	0.067	0.088	0.086	0.087	0.089	0.135
	S.D.	0.005	0.024	0.016	0.016	0.016	0.019	0.016	0.011
TFN	Mean				0.003	0.003	0.013	0.004	0.005
	S.D.				0.003	0.003	0.008	0.003	0.002

Note: The table reports the mean and standard deviations (S.D.) of the Monte Carlo sampling distributions (1,000 draws) of the estimates of TFE, TFA, TFN for all specifications of the baseline Table 1. In each Monte Carlo draw, the analysis is conducted on a random sample of ACLED events drawn with replacement.

Table B.11: Monte Carlo simulations testing ACLED event mismeasurement.

Probability of removal	0.01		0.1		0.2		0.5	
	TFA	TFE	TFA	TFE	TFA	TFE	TFA	TFE
Mean	-0.1139	0.0831	-0.1132	0.0836	-0.1109	0.0828	-0.1054	0.0813
S.D.	0.0009	0.0008	0.0028	0.0025	0.0044	0.0011	0.0063	0.0053

Note: The table reports the mean and standard deviations (S.D.) of the Monte Carlo sampling distributions (1,000 draws) of the baseline 2SLS estimates (col. 4, Table 1) of TFE and TFA for different probabilities of ACLED event mismeasurement for small groups. A group is defined to be small if  $(d^- + d^+) < 3$ , a condition that is satisfied for 49.7 percent of the groups in the sample.



Table B.12: Table multinomial logit.

CSF Surplus	0.000 (0.000)	
	Enmity link	Alliance link
Common enemy	-0.387** (0.183)	0.425** (0.181)
Common allied	0.097 (0.194)	-0.529*** (0.175)
Common allied and enemy	0.454*** (0.119)	-0.727*** (0.161)
Geodistance	-0.003*** (0.001)	-0.004*** (0.001)
Same ethnic group	0.842 (0.595)	0.865 (0.601)
Same Hutu Tutsi	0.831 (0.813)	0.550 (0.787)
Different Hutu Tutsi	1.295** (0.596)	-0.914 (1.148)
Zero Government	-1.977*** (0.711)	0.849 (0.763)
Zero Foreign	1.514*** (0.523)	0.761 (0.576)
Number of observations	9480	
Log likelihood	-445.92	

Note: Conditional logit estimator. The unit of observation is a pair of fighting groups in a given year. Alternatives correspond to Enmity (column 2), Alliance (column 3) and Neutrality as reference. Alternative-dependent group fixed effects are included. Standard errors in parentheses, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table B.13: RUM with time varying network.

	Dep. var.: Total Fighting		
	(1)	(2)	(3)
Total Fight. Enemies (TFE)	0.106*** (0.033)	0.114*** (0.035)	0.110*** (0.038)
Total Fight. Allies (TFA)	-0.112** (0.050)	-0.122** (0.053)	-0.137** (0.062)
Total Fight. Neutrals (TFN)	0.002 (0.006)	0.001 (0.006)	0.011 (0.023)
Kleibergen-Paap F-stat	4.0	3.6	5.1
Hansen J (p-value)	0.62	0.67	0.73
Observations	1040	1040	469
R-squared	0.554	0.579	0.597

Note: An observation is a given armed group in a given year. The panel contains 80 armed groups between 1998 and 2010. All regressions include group fixed effects and control for rainfall in the group's homeland. Robust standard errors corrected for Spatial HAC in parentheses. Column (1) replicates the baseline specification of the main table (column (4), Table 1) but with RUM. Column (2) replicates Column (1) of Table 3 but with RUM, while column (3) replicates column (4) of Table 3 but with RUM. Significance levels are indicated by \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

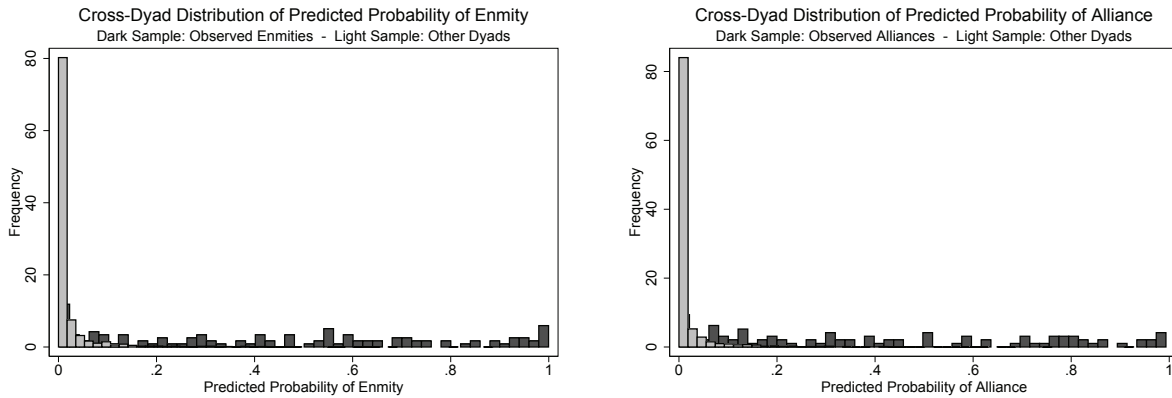
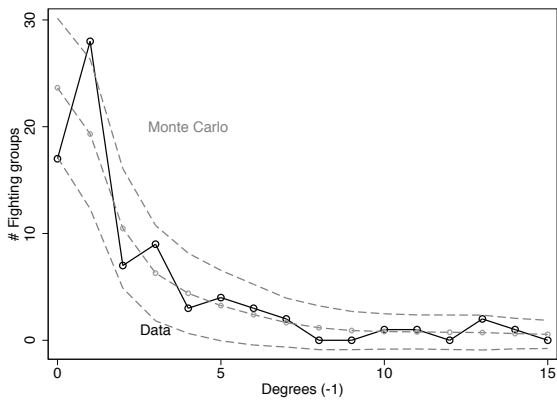
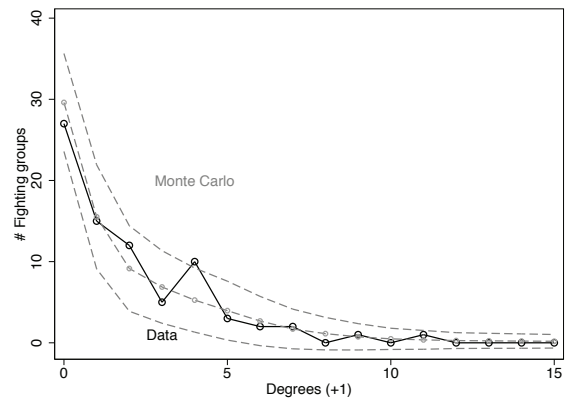


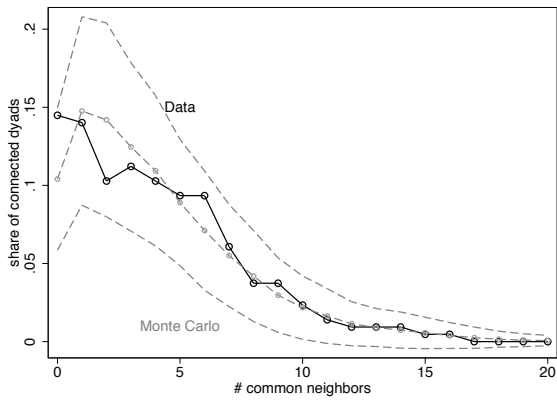
Figure B.1: *Predicted probabilities of enmities and alliances.*



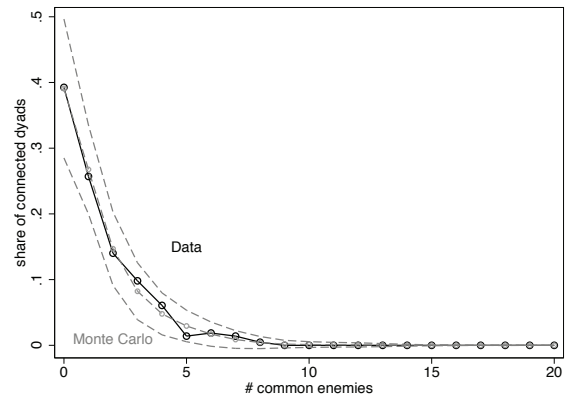
(a) Degree one enemies.



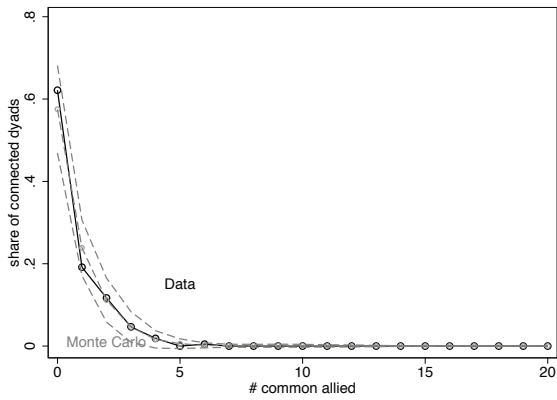
(b) Degree one allies.



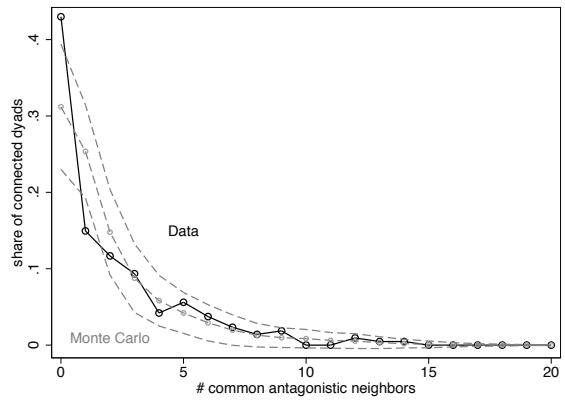
(c) Number of degree one links.



(d) Common enemies.



(e) Common allies.



(f) Conflicting neighbors.

Figure B.2: Goodness of fit statistics.

Table B.14: Welfare effect of removing armed groups with network recomposition.

Group	Sh. fight.	$-\Delta$ RD (exog. netw.)	$-\Delta$ RD (end. netw.)	Multipl. (end. netw.)	MAD	$-\Delta$ RD due to rewiring	New enm. at med.	New all. at med.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RCD-G	0.087	0.151	0.137	1.6	0.025	-0.014	1	-2
RCD-K	0.060	0.094	0.076	1.3	0.027	-0.018	2	0
Rwanda	0.053	0.066	0.103	1.9	0.040	0.037	-5	0
LRA	0.041	0.056	0.051	1.2	0.005	-0.005	0	-1
FDLR	0.066	0.055	0.058	0.9	0.008	0.004	-1	1
Mayi-Mayi	0.057	0.046	0.083	1.5	0.024	0.037	-2	1
Uganda	0.043	0.043	0.066	1.5	0.034	0.023	-4	4
CNDP	0.043	0.041	0.041	0.9	0.011	0.000	0	0
MLC	0.031	0.039	0.054	1.7	0.018	0.015	-2	1
UPC	0.022	0.030	0.011	0.5	0.020	-0.020	0	-1
Lendu Ethnic Mil.	0.024	0.022	0.049	2.0	0.020	0.027	-3	0
Mutiny FARDC	0.016	0.016	0.016	1.0	0	0	0	0
Interahamwe	0.014	0.014	0.027	2.0	0.024	0.013	0	-1
ADF	0.013	0.012	0.036	2.7	0.012	0.024	-2	1
FRPI	0.009	0.010	0.010	1.1	0	0	0	0

Note: The computation of the counterfactual equilibrium is based on the baseline point estimates of column 4 in Table 1. The results are based on 1,000 Monte Carlo simulations of an endogenous network recomposition. For each group, we report the observed share of total fighting involving this group (col. 1); the counterfactual reduction in rent dissipation associated with its removal (exogenous network) (col. 2); the counterfactual reduction in rent dissipation associated with its removal with network recomposition (col. 3); a multiplier defined as the ratio of col. 3 over col. 1 (col. 4); the Median Absolute Deviation in reduction in RD across Monte Carlo draws (col. 5); the difference between col. 3 and col. 2 (col. 6); post-rewiring number of new enmities and alliances at the median Monte Carlo draw (cols. 7-8).