THE GRAVITY OF VIOLENCE

Mathieu Couttenier[*] Julian Marcoux[†] Thierry Mayer[‡] Mathias Thoenig[§]

September 25, 2023

Preliminary and incomplete, please do not circulate

Abstract

This paper presents a framework for estimating and simulating a quantitative spatial model of trade and violence. In this new theoretical and empirical setup, suited for disciplining sub-national and international data, we first model the general equilibrium interactions between the economic and fighting margins in a micro-founded setup. We then show how the structural parameters can be recovered from the data in a simple and transparent way. A central element of the procedure consists in estimating a structural gravity equation of violence. Looking at sub-Saharan Africa over the period 1997 to 2022, we test the key predictions of the model and uncover new facts related to spatial frictions and conflicts. Finally, the model is used to quantify counterfactual policy interventions that aim at promoting development in weakly institutionalized contexts where insecurity is pervasive.

Keywords: Violence, conflicts, trade, gravity.

JEL Codes: F1.

[*]Department of Economics, ENS de Lyon; and CEPR.

[†]Department of Economics, University of Lausanne.

[‡]Department of Economics, Sciences Po, Paris; CEPII; and CEPR.

[§]Department of Economics, University of Lausanne; Oxford University; and CEPR.

1 Introduction

Even in contemporary times, armed conflicts and organized violence are pervasive phenomena that affect a significant portion of the world population. In the 2010s, approximately 12% of the world's population lived within conflict zones and 35% lived in countries experiencing a conflict, even though they were not directly exposed to violence (Korovkin and Makarin, 2021). This high prevalence of conflict carries significant implications for economic growth and development, particularly in low-income countries. The Organisation for Economic Co-operation and Development reported in 2009 that 60% of the world's poorest countries were affected by armed conflicts, which it considered a major hindrance to their ability to progress and develop (OECD, 2009).

Policies that work well in peaceful environments may not necessarily be effective in violent contexts. Recent evidence has shown that humanitarian aid and food assistance programs may inadvertently exacerbate violence in recipient countries (Nunn and Qian, 2014). Additionally, Crost et al. (2014) document instances in the Philippines where insurgents deliberately undermined a significant development program funded by the World Bank. Their motivation stemmed from the belief that the program's potential success could weaken their support base among the local population. These findings underscore the pressing need for a more comprehensive quantitative assessment of the functioning of development policies in conflict-prone countries. Recognizing this imperative, the United Nations has responded by adopting a Triple Nexus Development-Peace-Humanitarian doctrine aimed at addressing this concern.[1]

In order to design effective policy interventions in the most fragile regions, it is essential to comprehend the intricate general equilibrium interactions between economic factors and conflicts. However, a difficulty lies in the fact that the existing academic literature offers no data-fed quantitative framework that is capable of capturing these complex forces.[2] One example of this complexity can be observed in the ambiguous impact of road and infrastructure construction: on the one hand, it can enhance trade and appease tensions, while on the other hand, it may also facilitate border-crossing and intensify violence. Similarly, structural transformation can amplify rural-urban disparities, thereby rendering specific regions more vulnerable to conflicts.

The objective of this paper is to develop a quantitative spatial model that combines the economic and fighting margins in a tractable and estimable manner. Our approach is based on two premises. First, violence in Africa witnesses a notable involvement of groups with a remarkable ability for spatial projection. These groups extend the reach of their violence far beyond their rear bases, enabling us to perceive the violence in Africa as a flow with discernible origins and destinations. This perspective departs from the "place-based" analysis prevailing in existing empirical work on intrastate conflicts data and facilitates a more subtle understanding of the nature of violence in the African context. Second, in the tradition of the influential works of Olson (1993),

[1] `https://www.un.org/humansecurity/wp-content/uploads/2022/03/FINAL-Triple-Nexus-Guidance-Note-for-web_compressed.pdf`

[2] Quoting McGuirk and Burke (2020a) "Modeling general equilibrium forces is important as economic shocks that alter the opportunity cost of violence could also affect the spoils of victory or a government's capacity to repel insurgents, yielding an unclear relationship. This ambiguity is reflected in a markedly inconclusive empirical literature, characterized by inconsistent findings and by significant identification challenges: income may affect conflict; conflict may affect income; and both".

Tilly (1985) and Dal Bó and Dal Bó (2011), we model conflict and war-making as organized crime: there, fighting groups can be primarily seen as (stationary or roving) bandits exerting violence to appropriate population's income. This view has received considerable empirical support in the literature.[3]

We start by building our general equilibrium theoretical framework. Then, we show how the structural parameters can be recovered from the data in a simple and transparent way. A central element involves estimating a theory-consistent gravity equation of bilateral flows of violence. Furthermore, looking at sub-Saharan Africa over the period 1997 to 2022, we test the key predictions of the model and uncover new facts related to spatial frictions and conflicts. The overall approach is portable and frugal in terms of data requirements, which proves advantageous when analyzing conflict zones characterized by economic deprivation and limited data availability. Finally, the model is used to quantify counterfactual policy interventions that aim at promoting development in weakly institutionalized contexts where insecurity is pervasive. Specifically, we evaluate the Great Lakes Initiative, a recent trade facilitation policy designed by the World Bank to improve border-crossing infrastructures between DRC, Rwanda and Uganda. At the aggregate level, our simulations show that the policy would contribute to increasing economic opportunities and pacifying the Great Lakes region. Quantitatively, we find that the welfare gains attached to trade facilitation increase by twenty percent when the general equilibrium feedback effect on violence and conflict is factored in the policy simulation.

A modeling challenge relates to the coupling of the trade and fighting decision margins. We make progress thanks to the (overlooked) observation that trade and conflict models share a fundamental conceptual commonality. They can both be grounded in random utility discrete choice models that generate CES-like functional forms governing aggregate behaviors (referred to as the Tullock Contest Success Function, CSF, in conflict models). This implies that we can use recent advances in methods proposed by the quantitative spatial economics literature to model the conflict margin (Redding and Rossi-Hansberg, 2017; Allen et al., 2020). Of particular importance is our derivation of a "structural gravity" of violence, very similar to the trade-in-goods counterpart, a major building block for quantitative spatial models. Our main innovations reside in deriving foundations for and empirically estimating a gravity equation of violence. This equation turns out to be essential for recovering spatial frictions, themselves crucial to the overall quantification of the model. This approach allows for a tractable connection between fighting and economic equilibria.

[3] Scholars working on conflicts typically assign the causes of conflict to two categories: greed and grievances (Collier and Hoeffler, 2004). It is now well-documented empirically that greed emerges as a primary driver of violence. The direct form of appropriation encompasses various activities such as looting, extortion, forced labor, and the outright theft of land and resources. This form of appropriation is particularly evident in the control exerted over natural resources like oil and minerals, where competition for resource control becomes a catalyst for conflict (Dube and Vargas, 2013; Berman et al., 2017; Sanchez de la Sierra, 2020). Furthermore, conflicts can be fueled by land disputes arising between transhumant pastoralists and sedentary agriculturalists, which stem from issues related to land tenure and unequal access to fertile land (McGuirk and Nunn, 2020; Eberle et al., 2020; Berman et al., 2021). The lens of appropriation sheds light on numerous other causes of conflicts as well. Political, cultural, and religious motives are frequently employed as justifications by armed groups; however, they can also serve as a smokescreen to mask acts of appropriation. This dynamic is particularly noticeable in secessionist movements and ethnic conflicts, where the pursuit of resource control and territorial dominance often underlies the surface motivations.

Literature: The literature on conflicts has experienced a notable surge in empirical advances over the past fifteen years, an evolution notably driven by the availability of geolocalized data on violent events at a fine-grained level. Noteworthy contributions have focused on exploring the influence of local factors in shaping the geographical patterns of violence. These factors include economic inequality (Buhaug et al., 2011), distance to borders and capitals (Buhaug and Rød, 2006) and political exclusion (Cederman et al., 2009). More recent studies have further made use of this high-resolution information to investigate the causal relationship between exogenous micro-level economic shocks and the likelihood of local conflicts (Dube and Vargas, 2013; Berman et al., 2017; Harari and Ferrara, 2018; McGuirk and Burke, 2020a). However, these studies primarily employ reduced-form empirical models and, although many of them provide insightful theoretical arguments, they do not explicitly aim to bring formal models to the data.[4] As a consequence, they are not equipped for quantifying the welfare implications of counterfactual policy interventions. Henceforth, their ability to assess the precise impact of policy measures remains limited.

Aside from the seminal theoretical contribution by Dal Bó and Dal Bó (2011), the academic literature has largely overlooked the examination of general equilibrium interactions between conflict and economic activities. Despite repeated calls from prominent scholars for a better understanding of these mechanisms (Dell et al., 2014; Burke et al., 2015), there remains a gap in research on this topic. A limited number of structural studies have explored the theoretical and empirical aspects of violence diffusion in spatial networks of fighting groups (König et al., 2017; Amarasinghe et al., 2020; Mueller et al., 2022). These attempts provide valuable predictions related to the violent behavior of geographically located fighting groups. However, they adopt a partial equilibrium analysis that captures the impact of economics on conflict in a reduced form, typically represented by an exogenous "prize of the contest". As a consequence, they are silent on the feedback loop from the fighting equilibrium on the economy, thus leaving various channels through which economic activities and violence may potentially interact unexplored.

Finally, our paper contributes to the modern literature on spatial economics by extending the existing framework of Quantitative Spatial Models (QSM) to address the context of weakly institutionalized countries. Doing so, we deviate from the standard structure of QSMs by accounting for a key characteristic of these countries, namely the presence of a violent context characterized by inadequate enforcement of property rights and limited state capacity. Our research aligns with the recommendation put forth by Proost and Thisse (2019) in their survey of the spatial economics literature: "when the economy is mainly informal and market institutions do not function well, we need new models".

The remainder of the paper is as follows: Section 2 sets up the theoretical microfoundations of our framework, and shows how the two main gravity equations for trade and violence are derived and also how they interact in the general equilibrium of our model. The gravity equation of bilateral violence is the main empirical novelty of our paper, and is estimated in Section 3. Section 4 then takes estimated frictions from this estimation and explains how the rest of parameters needed

[4] Some of these papers (Berman et al., 2017, for instance) estimate theory-free models that examine the spatial decay of predictors of violent events, such as the impact of a spike in the global market price of locally produced resources (e.g., minerals or crops) on violence at various distances (e.g., 50km, 100km, 500km).

to simulate the model are recovered from combining the structure of the model with observables. Section 5 runs counterfactual simulations of relevant policy changes in Africa to analyze the consequences of a trade facilitation experiment in a framework explicitly accounting for violence.

2 Theory

Our theoretical setup is modeling the interactions between two types of activities: production of a tradable good and appropriation of that good. Individuals in each location of a multi-regional world choose in which activity to allocate their labor force: they adopt the status of worker or fighter. For each activity, there are inter-regional frictions that hamper both the ability of workers to ship their production on non-local markets, and the ability of fighters to steal income generated at long distance from their home base. Those frictions will generate two gravity-like equations, one for trade in goods, one for bilateral flows of violent activity. Those two equations are not independent, being linked by labor market clearing conditions and optimal occupational choice in each region to yield general equilibrium. We start with the description of the two gravity equations, and then turn to equilibrium.

2.1 Gravity of goods

There are N regions indexed by i, with a population of $\bar{L}_i$ that freely allocates between farming (L_i) and fighting (l_i). Each region i produces a single variety of the farming (tradable) good, and is the sole source of this variety. Consumers in region n have a CES ($\sigma > 1$) utility over all available varieties, given by $U_n = \left(\sum_i (q_{in})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$. We further assume perfect competition and iceberg trade costs τ_{in}, the market price being $p_{in} = \tau_{in} w_i^P / A_i$, with w_i^P and A_i representing the wage of farmers (producers) and productivity, respectively. The share of aggregate expenditure E_n that consumers in region n spend on the variety from i is given by:

$$\pi_{in} \equiv \frac{(\tau_{in} w_i^P / A_i)^{1-\sigma}}{\sum_k (\tau_{kn} w_k^P / A_k)^{1-\sigma}}. \tag{1}$$

The equilibrium bilateral trade flows is given by:

$$Y_{in} = \pi_{in} E_n = \tau_{in}^{1-\sigma} \times \left(\frac{w_i^P}{A_i}\right)^{1-\sigma} \times \frac{E_n}{\sum_k \left(\frac{\tau_{kn} w_k^P}{A_k}\right)^{1-\sigma}}, \tag{2}$$

which is the classical formulation of the gravity equation for goods.[5]

The aggregate trade revenues of producing region i (including internal trade) are the sum of

[5] Note that many microfoundations lead to the same aggregate bilateral trade equation, including Armington (1969), Anderson and van Wincoop (2003), Krugman (1980), Anderson et al. (1992), Eaton and Kortum (2002), and Chaney (2008) models (Head and Mayer, 2014, provide an overview of those foundations).

trade revenues from all destination regions n:

$$w_i^P \times L_i = \sum_n \pi_{in} E_n, \tag{3}$$

where L_i is the labor supply of farmers in region i. Together, these equations provide the framework for understanding the partial equilibrium of trade in our model.

2.2 Gravity of violence

Let us now turn to patterns of inter and intra-regional violence, viewing violence as an appropriation game. In this game, farmers and fighters compete for a share of the total income in each region. The game is played over an infinite horizon with repeated contests, and each stage of the game involves a share s_n of each player's income being definitively secured as an absorbing state. In equilibrium, income ends up being ε-secured, where $\varepsilon > 0$ represents the share of unsecured income. The fighting revenues obtained are repatriated to the origin region i, where production, trade, and consumption take place.[6]

An important distinction in the model is between aggregate gross income Y_n (secured and unsecured) and aggregate expenditure E_n (fully secured). With the assumed micro-foundations of our model, the aggregate ratio E_n/Y_n is not affected by the endogenous composition of farming and fighting:

$$E_n = s_n Y_n. \tag{4}$$

Each destination region n is characterized by the unsecured share $(1 - s_n)$ of income that is looted by fighting groups. We interpret s_n as an *exogeneous* state capacity. We assume sequential periods where all agents (farmers and fighters) first receive their wages in n (front-loaded payment); then looting of money takes place and money is repatriated in the origin region i and finally agents produce and consume in their origin region. The origin region i hosts one fighting group that recruits l_i fighters at a local competitive wage w_i^F. The optimal assignment of fighters l_{in} to different regional "targets" n is subject to a spatial friction factor ξ_{in}, which allows for local looting ($\xi_{ii} < \infty$), raiding ($\xi_{in} < \infty$ with $i \neq n$), and no military action ($\xi_{in} = \infty$).

On the battlefield of region n, each group produces observable violence with constant returns to scale (CRS) technology, denoted as $\texttt{violence}_{in} = \psi_i l_{in}$. The operational performance of group i is calculated as the ratio of violence produced by group i to the spatial friction factor, multiplied by a Frechet unobservable shock, $\tilde{u}_{in}$, with shape parameter $\gamma > 0$ which captures how homogeneous is military luck over different military group-destination combinations. Victory in the battle for lootable income in n goes to the group with largest operational performance. From our functional

[6]To account for the looting of resources, we need to perform an accounting of the looted resources at each stage of the game. Farmers are looted by fighters, who are in turn looted by other fighters, and so on. The model assumes a front-loaded payment of farmers as unsecured income, and fighters are allocated across regions n. The end game period T has an order of magnitude given by $\max_n (1 - s_n)^T < \varepsilon$. As a result, the model allows for asymptotic convergence only for $\varepsilon = 0$. See Appendix A for details.

form assumptions, the probability that i succeeds is:

$$p_{in} \equiv \text{Prob}\left(\frac{\psi_i l_{in} \tilde{u}_{in}}{\xi_{in}} > \frac{\psi_k l_{kn} \tilde{u}_{kn}}{\xi_{kn}}, \quad \forall k \neq i\right) = \frac{(\psi_i l_{in}/\xi_{in})^\gamma}{\sum_k (\psi_k l_{kn}/\xi_{kn})^\gamma}. \tag{5}$$

This equation describes the probability of success and therefore the incentives for a fighting group in i to launch operations in n. This is a key element of the gravity of violence. The second brick required to obtain gravity relates to the spatial allocation of troops. The optimal allocation of troops across regions n (l_{in}) aims to maximize gross fighting revenues. For tractability, we assume atomistic players. A first advantage is to simplify interactions such as there is no oligopoly on the appropriation market; here, each player best-responds to other players' actions but each player does not internalize her impact on aggregate violence. Second, the continuum assumption implies that the winning probability p_{in} can be interpreted as a realized share of income in n captured by fighters from i.[7] Expected revenues of fighters in i therefore match its realized value, $\mathbb{E}(R_i) = R_i$, and the objective function is given by:

$$R_i \equiv \max_{\{l_{in}\}} \sum_n p_{in} \times (1 - s_n) Y_n \qquad \text{s.t.} \quad l_i = \sum_n l_{in}.$$

In order to obtain a concave maximization problem of R_i with respect to all optimal allocations of fighters l_{in} by group i, we need $0 < \gamma < 1$ (which we estimate to be the case in empirics). This will also ensure that an interior solution exists. Solving for the optimal allocation of troops and using the resulting l_{in} across regions in (5) yields the partial equilibrium flow of violence (quantity) from i to n:[8]

$$\texttt{violence}_{in} \equiv \psi_i l_{in} = \xi_{in}^{-\frac{\gamma}{1-\gamma}} \times \left(\frac{\psi_i}{w_i^F}\right)^{\frac{1}{1-\gamma}} \times \frac{(1 - s_n) Y_n}{\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_k}{w_k^F}\right)^{\frac{\gamma}{1-\gamma}}}. \tag{6}$$

This equation for the equilibrium flow of violence resembles gravity equations in trade/migration models (one can see the parallel by comparing with equation 2). Economic shocks such as changes in wages impact violence in complex ways, channeled by the full spatial structure of the model. An increase in opportunity cost (w_i^F) leads to a decrease in violence emanating from i, while an increase in income at destination (w_n^P, and therefore Y_n) results in an increase in violence. Those mechanisms are intuitive. A new effect, captured by the $\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_k}{w_k^F}\right)^{\frac{\gamma}{1-\gamma}}$ term, emerges as a result of increased competition among groups from different regions fighting for income in n. A rise in w_k^F in any region k leads to an increase in violence from i to n through decreased competition on the battlefield and therefore a higher incentives for i to send fighters in n. We refer to this effect as *the Multilateral Resistance of Violence*, MRV, in parallel to the terminology used in trade (Anderson and van Wincoop, 2003, were the first to use this terminology) in order to account for competition in a given destination market (the denominator term in equation 2). As for gravity in goods regressions, omitting multilateral resistance terms (which accounts for spatial interdependence) is a source of

[7] This modeling strategy is similar to various approaches in trade which micro-found aggregate gravity flows out of individual behavior of heterogeneous agents (see section 2.3 of Head and Mayer (2014)).

[8] The full derivation of equilibrium flow of violence is given in Appendix B.

mis-specification and bias in the main variables of interest, here frictions ξ_{in}. Note that equation (6) is a "quantity" equation for violence, which is what we observe in the dataset at hand. The equivalent of (2), describing the bilateral financial flows linked to violence, is $\texttt{violence}_{in} \times \left(\frac{w_i^F}{\psi_i}\right) = w_i^F l_{in}$. Last, we can write an expression for the gross income accruing to fighters in region i:

$$R_i = w_i^F \times l_i = w_i^F \times \sum_n l_{in}. \tag{7}$$

2.3 General equilibrium

General equilibrium: In equilibrium, free occupational choice between farming and fighting ensures equalization of (fully secured) incomes, and thus wages in i:

$$s_i w_i^P = s_i w_i^F = s_i w_i. \tag{8}$$

Gross Nominal Income can therefore we rewritten as

$$Y_i = w_i^P L_i + w_i^F l_i = w_i \bar{L}_i, \tag{9}$$

with the last equality ensured by the labor market clearing condition under which fighters and farmers add up to total workforce $\bar{L}_i$:

$$\bar{L}_i = L_i + l_i$$

Two parallel additional market clearing conditions must hold in equilibrium. The first one is that the sum of demands for the good produced in i has to add up to production in n. This is equation (3), which—combined with (2) and accounting for free occupational choice—can be rewritten as

$$w_i L_i = \sum_n \frac{\tau_{in}^{-(\sigma-1)} \left(\frac{A_i}{w_i}\right)^{\sigma-1}}{\sum_k \tau_{kn}^{-(\sigma-1)} \left(\frac{A_k}{w_k}\right)^{\sigma-1}} s_n w_n \bar{L}_n. \tag{10}$$

The second market clearing condition is that the bilateral flows of revenues obtained from appropriation sum to total revenues of violence in i (equations 6 and 7):

$$w_i l_i = \sum_n w_i l_{in} = \sum_n \frac{\xi_{in}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_i}{w_i}\right)^{\frac{\gamma}{1-\gamma}}}{\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_k}{w_k}\right)^{\frac{\gamma}{1-\gamma}}} (1 - s_n) w_n \bar{L}_n. \tag{11}$$

Existence and uniqueness: In order to establish the existence and uniqueness of general equilibrium, it is useful to write a fixed-point "master equation" system that combines the fighting and trade revenue equations. Adding up (10) and (11), one obtains

$$w_i \bar{L}_i = \sum_n \beta_{in}(\mathbf{w}) w_n \bar{L}_n, \tag{12}$$

where $\beta_{in}(.)$ are non-linear functions of the wage vector $\mathbf{w}$:

$$\beta_{in}(\mathbf{w}) \equiv (1-s_n) \times \frac{w_i^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_i}{\xi_{in}}\right)^{\frac{\gamma}{1-\gamma}}}{\sum_k \left(\frac{\psi_k}{\xi_{kn} w_k}\right)^{\frac{\gamma}{1-\gamma}}} + s_n \times \frac{w_i^{1-\sigma} \left(\frac{A_i}{\tau_{in}}\right)^{\sigma-1}}{\sum_k \left(\frac{A_k}{\tau_{kn} w_k}\right)^{\sigma-1}}. \tag{13}$$

In order to prove existence and uniqueness of a general equilibrium in the vector of wages, we apply in the appendix resolution techniques from spatial economics. Specifically, we define the excess demand function and apply Alvarez and Lucas (2007) and Mas-Colell et al. (1995) to establish the desired properties. The isomorphism between the trade and fighting equations allows us to apply these techniques and conclude that a general equilibrium exists and is unique in our model.

3 Empirical Gravity of Violence

3.1 Econometric Methodology

Based on our theoretical prediction, we estimate a gravity equation to reveal empirical determinants of bilateral violence. The equation is expressed in terms of the share of violence at destination n. Allowing for time variation (indexed t) in equation (6), we obtain:

$$\frac{\texttt{violence}_{int}}{\texttt{violence}_{nt}} = \frac{\psi_{it} l_{int}}{\sum_k \psi_{kt} l_{knt}} = \xi_{in}^{-\frac{\gamma}{1-\gamma}} \times \left(\frac{\psi_{it}}{w_{it}}\right)^{\frac{1}{1-\gamma}} \times \left[\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_{kt}}{w_{kt}}\right)^{\frac{1}{1-\gamma}}\right]^{-1}. \tag{14}$$

We follow the now standard practice in the gravity literature to estimate this equation using Poisson pseudo-maximum likelihood (PPML) with high-dimensional fixed effects[9], translating into:

$$\mathbb{E}\left(\frac{\texttt{violence}_{int}}{\texttt{violence}_{nt}}\right) = \exp\left\{-\frac{\gamma}{1-\gamma}\log \xi_{in} + \texttt{FE}^o_{it} + \texttt{FE}^d_{nt}\right\}, \tag{15}$$

where $\texttt{FE}^o_{it}$ and $\texttt{FE}^d_{nt}$ are origin×year and destination×year fixed effects, respectively. The estimation provides spatial frictions and fixed effects, which are crucial objects for the next step when we turn to the task of recovering structural parameters of the model.

3.2 Data Sources

Conflict data: We use conflict event data from the Armed Conflict Location and Event Dataset (ACLED) which contains information on conflict events in all African countries from 1997 to 2022 (`https://acleddata.com/`).[10] Crucially, these data contain information about the date, GPS location, nature of events (including the list of types), as well as who are the actors that participate to each single event. The data are widely used (Berman et al., 2017; McGuirk and Burke, 2020b;

[9] Fally (2012) showed PPML to be a natural estimator for micro-founded structural models of gravity, allowing for a structural interpretation of the fixed effects as well as a large share of zeroes.

[10] Download on June 1st, 2022.

Moscona et al., 2020; Berman et al., 2021). Events are compiled from various sources, including press accounts from regional and local news, humanitarian agencies, and research publications.

The dataset contains information on 487,168 event-actor pairs (observations), involving 6,878 actors over 281,311 distinct violent events, as multiple actors may participate in a single event. Those actors are grouped into eight different "types". Our two types of interest are "Rebel groups" and "Political militias", which are the most relevant for our purpose since they are the most likely to be involved in the violence-for-appropriation acts that we are modeling. They are also the ones most susceptible to export violence outside their home base, and the ones for which we are most likely to be able to locate a home base. "Identity Militias" are another type that could be relevant for our analysis, but most of their violence is reported by ACLED to be "communal violence" and they are probably too small for their rear base to be documented (Table 1 reports that those are a large share of groups but that they are present in only 8.4% of violent events). Together, rebels and political militias represent more than a quarter of actors, are involved in about half the events, and nearly a third of observations.

Table 1: Actors of violence in ACLED

Type of actors	% Events	% Groups	% Obs.
Rebel groups	24.6	5.4	14.4
Political Militias	25.2	20.2	15.2
State Forces	36.5	13	21.8
Identity Militias	8.4	55.9	6.2
Rioters	11	.9	6.9
Protesters	23.8	1	13.7
Civilians	32.7	1.4	18.9
External/Other Forces	4.8	4.9	2.8

Note: The first column reports the percentage of events in which at least one member of the type of actor is involved. The second column gives the percent of distinct group in ACLED, and the third column the percentage of total event-actor observations.

The second row of Table 2 shows that restricting attention to those two groups reduces drastically the sample in terms of the distinct number of violent groups (1,751) and events (137,518). Note that events that involve for instance a rebel group attacking civilians is not dropped. We simply drop the row of the dataset reporting the civilians (victim of violence in that case) for that event (which retains information on the precise location where the civilians were attacked).

To ensure high geographical precision of our sample, we exclude events that are coded as "part of a region", "region", or "country" (based on the variable "geo" taking the value of 2 or 3 in the original dataset), leaving us with 93,137 distinct events. ACLED also documents the type of each event. We remove events coded as "agreements" or "other". After a final filter merging groups with similar names, our sample totals 63,151 events involving 1,335 actors (last column of Table 2). The detail of how violent events are distributed across event types is provided in Table 3. The categories of events related to battles, violence against civilians, explosions, and remote violence represent over 90 % of our final dataset. A reassuring finding in this table is that protests and riots which are probably the events that depart the most from our theoretical setup represent over a

Table 2: From raw data to our sample

Steps	# Events	# Groups	# Obs.
0.Raw	281311	6878	487168
1- Rebels groups + pol. militias	137518	1751	144347
2a- Geographic filter #1	132673	1681	139192
2b- Geographic filter #2	93137	1458	97730
3- Duplicates	92880	1458	96616
4- Drop Sub-events: Other & Sub-events	91853	1409	95533
5a- Name filter #1	91853	1391	97882
5b- Name filter #2	63151	1335	68439

Note: ACLED reports a spatial precision codes based on the source material notes. *"Geographic filter #1"'* excludes events where source material notes that events took place in a region. *"Geographic filter #2"* excludes events where source material notes that events took place in a small part of a region. *"Step 5a- Name filter #1"* is made up of various changes designed essentially to make group names consistent over the period, rename misspelled group names, associate militia to the country from which they originate, and to associate factions with the main groups from which they originate. *"Step 5b- Name filter #2"* consists of excluding events where the actor is unidentified.

quarter of violence in the original data, while they are reduced to 0.5% by our filters.

Table 3: From raw data to our sample

Steps	Step 0-Raw	Step 5b- Name filter #2
Type of violence	% obs	% obs
Battles	30	54.2
Explosions/Remote violence	8.5	10.2
Protests	15.6	.2
Riots	11	.3
Strategic developments	6.4	7.9
Violence against civilians	28.5	27.2

Note: ACLED categorizes events into the following types: *Battles* correspond to violent confrontations between two politically organized armed groups; *Explosions/Remote Violence* refer to events involving one-sided acts of violence where the aggressor uses tools or tactics that prevent the targeted group from effectively responding; *Protests* are nonviolent public demonstrations where participants may face violence from others; *Riots* correspond to events involving acts of violence and disruption carried out by demonstrators or mobs; *Strategic Developments* encompasse events that may not directly involve political violence, but they can potentially trigger future violent events; *Violence Against Civilians* refer to deliberate acts of violence inflicted by organized armed groups upon unarmed non-combatants.

Other: For the purpose of recovering the structural parameters of the model, we make use of various databases. Information on nighttime light data is obtained from the harmonized global nighttime light dataset (Li and Zhou, 2017). Population data is obtained from WorldPop using the top-down unconstrained estimation modeling approach for 2000. We identify the main crop(s) produced by each cell using data from the FAO's Global Agro-Ecological Zones. This dataset is constructed from models that use location (climate information and soil characteristics) and crops' characteristics to generate a global GIS raster of the suitability of a grid cell for cultivating each crop (17 crops for which we have a worldwide prices). Global crop prices (base 100 in 2000) are obtained from the World Bank Commodity Dataset (World Bank Group, 2020).[11]

3.3 Vectorization of Georeferenced Conflict Events

Here we describe our procedure of "vectorization" that process events into bilateral flows of violence with a magnitude and a direction ($\texttt{violence}_{int}$). In the model, each fighting group g is associated with a rear base $rb(g) \in i$, where recruitment takes place and fighters are paid the prevailing competitive wage of the local labor market in the region of origin i. The bilateral flow of violence between i and n is given by:

$$\texttt{violence}_{in} = \sum_{g} \mathbb{I}_{rb(g)\in i} \quad \times \quad \sum_{t} \texttt{\#events}_{gnt}, \tag{16}$$

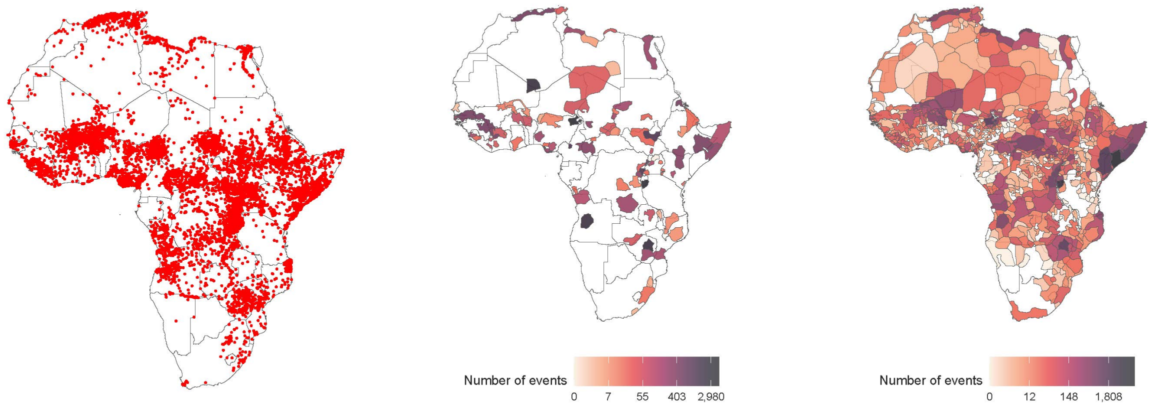
where $\texttt{\#events}_{gnt}$ is the number of events attributed to group g in n over time period t, and $\mathbb{I}_{rb(g)\in i}$ is a dummy variable which takes a value of 1 if the rear base of the group g is located within region i. This defines a cross-section of events between geographical units i and n as our dependent variable.

We need to define and measure $rb(g)$, the location of the rear base of armed groups. In the absence of exhaustive information on the actual rear base of armed groups at the scale of the African continent, we assume in most of our analysis that an armed group's credible rear base stems from its ties to ethnic groups. With the aid of Murdock's data (Murdock, 1959), which allow us to identify historical ethnic homelands, we can allocate the armed group's rear base based on its ethnic affiliation. Hand-collected information has been gathered to match 182 armed groups with 81 ethnic homelands (over the 824 ethnic regions defined in Murdoch). The collection of this information has been done on the largest armed groups, i.e. on the 373 groups that have participated at least to three events and fight at least over three years. They represent 62,577 event-actor pairs. Thus, for all ACLED events involving an armed group for which we are able to find ethnic ties, vectorization can be performed by designating an origin i and destination n for the event (both being ethnic homelands). In the end, we found information for ethnic ties of 182 groups , we compute the total number of events for each origin-destination pair from 1997 to 2022, leaving us with 81 origin of violence and 486 destinations of violence (mapped in figure 1), for a total of 39,366 dyads. Out of this universe of potential flows, 97% of them are zero-violence flows.

As a first motivation for our specification of frictions, Figure 2 presents evidence of unconditional correlation between (log-)distance and the share of events by destination (our dependent variable in the regressions). In the overall sample, the binscatter in panel (a) shows a strong negative correlation (explaining 73% of the variance of the share of events), which is maintained in panel (b) when we exclude internal violence, i.e., violence that originates and ends within the same location. On the contrary, we observe that the internal distance within an ethnic group has virtually no impact on the share of events. This suggests that one needs to account for the effects of ethnic

[11]Night time light data: `https://www.mdpi.com/2072-4292/9/6/637`; World Pop: `https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained/`; FAO's Global Agro-Ecological Zones: `https://gaez.fao.org/`

Figure 1: Data visualization



Note: The left panel displays the set of geolocalized ACLED events used to build the bilateral flows of violence. The middle panel displays the 81 ethnic groups which observed violence is originating from. The right panel displays the 486 ethnic groups that experience violence on their territory.

border-crossing (and country border-crossing) on top of the simple effect of distance, in order to account for the distinct friction suggested by the distinct patterns in panels (b) and (c) of figure 2.[12]

3.4 Econometric Specification

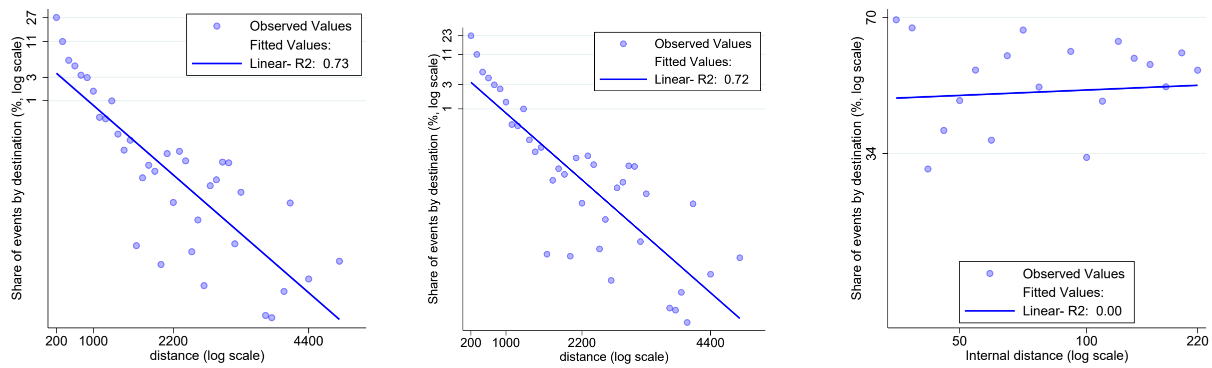
The empirical gravity of violence equation (15) is estimated using a specification where ξ_{in} includes three observable frictions: bilateral distance, ethnic homeland border-crossing, and country border-crossing. The structural error term comes from an unobservable bilateral impediment to violence, denoted ν_{in}. The dependent variable is the share of violence in n which comes from i. The independent variables are the log of the bilateral distance between i and n, a binary variable that equals one if i and n share the same ethnic homeland, and another binary variable that equals one if i and n are in different countries. None of those friction variables having a time dimension, we sum the events over the whole period, to obtain a cross-section of 81 origins and 486 destinations, for a total of 39,366 dyads. Estimation is carried out using the HDFE PPML estimator (which also allows to handle the very large proportion of zeroes), including origin and destination fixed effects, and clustering by dyad:

$$\mathbb{E}\left(\frac{\texttt{violence}_{in}}{\texttt{violence}_{n}}\right) = \exp\left[-\frac{\alpha_1\gamma}{1-\gamma}\log\texttt{dist}_{in} - \frac{\alpha_2\gamma}{1-\gamma}\texttt{ethnic}_{in} - \frac{\alpha_3\gamma}{1-\gamma}\texttt{border}_{in} + \tilde{\nu}_{in} + \texttt{FE}_i^o + \texttt{FE}_n^d\right], \quad (17)$$

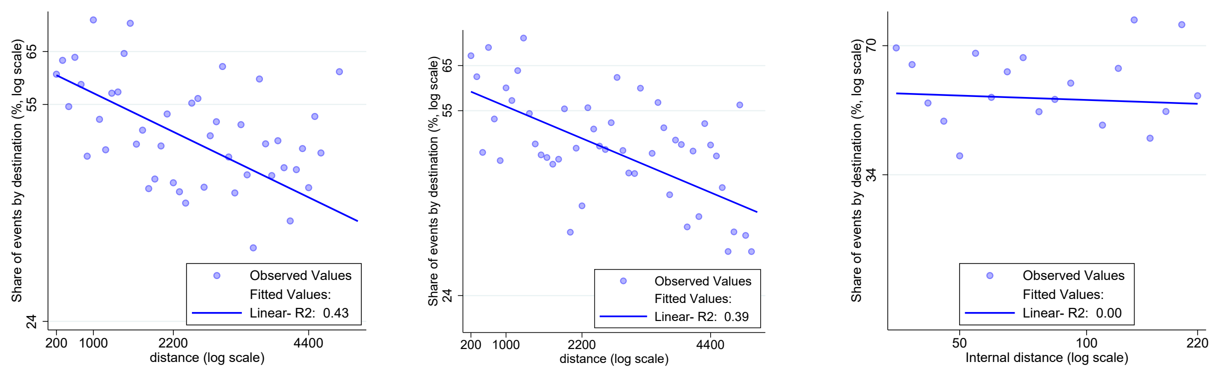
where $\tilde{\nu}_{in} \equiv -\frac{\gamma}{1-\gamma}\nu_{in}$.

[12]Panels (d) to (f) show very similar patterns for positive flows.

Figure 2: Unconditional correlation between distance and share of events by destination



(a) Full sample (b) Outside ethnic homeland (c) Inside ethnic homeland



(d) Full (shares > 0) (e) Outside (shares > 0) (f) Inside (shares > 0)

Note: the graphs represent several binscatter plots. Panels (a), (b), (d) and (e) display 50 bins, whereas panels (c) and (f) display 20 bins. Panels (a), (b), (c) consider all flows of violence, whereas panels (d), (e) and (f) are restricted to positive flows.

3.5 Results

Table 4 displays the estimates. The elasticity of violence with respect to distance is estimated to be -2.84 (column 1). This implies that doubling the distance between two ethnic groups is associated with a 86% ($\exp(-2.84 \times \log 2) - 1$) drop in the share of events from the ethnic group of origin. To put this elasticity in context, it is worth comparing it to the elasticity of trade with respect to distance. Head and Mayer (2014) report that the elasticity of trade to distance (not distinguishing by transport mode) is -1.1 on average in a meta-analysis of 328 estimates. The elasticity of trade using land transport (mostly trucks in advanced economies) is estimated to be around -2 by Combes et al. (2005) This suggests that the logistics of violence in Africa, which heavily relies on road transportation, may be a significant factor driving the observed spatial patterns of violence. From columns (2) to (4), we sequentially add variables capturing the frictions associated with crossing an ethnic border or a political border. Crossing ethnic borders increases the likelihood of violence by nearly an order of magnitude ($\exp(2.28) = 9.8$), which suggests that raiding other ethnic groups to capture their income is a key driver of violence. On the other hand, crossing political borders

Table 4: Estimates of the frictions

	(1)	(2)	(3)	(4)
(log-)distance	–2.840***	–3.107***	–2.700***	–2.679***
	(0.069)	(0.074)	(0.077)	(0.077)
Ethnic border		2.277***	2.353***	2.271***
		(0.260)	(0.263)	(0.268)
Pol. border			–1.399***	–1.765***
			(0.150)	(0.248)
Pol. border $\times$ one split				0.480*
				(0.267)
Pol. border $\times$ $\geq$2 splits				0.732**
				(0.351)
Observations	39,366	39,366	39,366	39,366

Note: Dependent variable is the share of events in *n* originating in *i*. Cross-section over 1997-2022. Estimated with PPML with *i* and *n* fixed effects. * Significant at 10%, ** significant at 5%, *** significant at 1%.

leads to a three quarters drop in violence ($\exp(-1.4) - 1 = -0.75$) , indicating that national border security measures can be effective in reducing violence. Additionally, in column (4), we account for the fact that ethnic groups might be crossed by one or more national borders. When the ethnic group of origin is divided between at least 3 countries, crossing political borders "only" results in a 42% ($\exp(-1.76 + 0.48 + 0.732) - 1$) drop in violence. This result suggests a coordination problem since the violence-reducing impact of national borders is lower when several governments have to manage local violence. Overall, these findings highlight the importance of considering both ethnic and political factors in understanding violence, and suggest that policies targeting border security and raiding could be effective in reducing violence in conflict-affected regions.

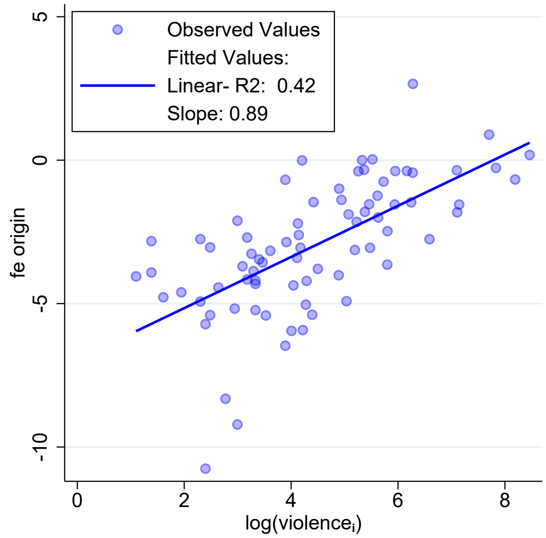
4 Model Inversion

In this section, we present the procedure that we use to recover all structural parameters of the model needed for the counterfactuals run in the next section. The procedure involves six steps.

Step 1: We estimate the gravity equation of violence (equation 17) to obtain coefficients on the three frictions and on the origin fixed effects $\widehat{\text{FE}}_i^o$. One origin, labeled 1, must be taken as a reference,[13] such that $\widehat{\text{FE}}_1^o = 0$. In order to be able to meaningfully compare the fixed effects, we restrict the regression to the largest connected set, which comprises 77 origins. In Figure 3, we display the (expected) positive correlation between $\widehat{\text{FE}}_i^o$ and the (logged) number of violent events emanating from *i*.

[13]The reference group is chosen as being the Zaghawa group located in between Sudan and Chad. The armed group *Justice and Equality Movement* has ethnic ties with the Zaghawa group.

Figure 3: Correlation Fixed Effects of Origin and the (log-) level of violence



Note: The estimated origin fixed effects on the y-axis is obtained from the estimating equation (17), with sample restricted to the largest connected set and friction coefficients reported in column 3 of Table 4.

Step 2: Using the theoretical correspondence of the origin fixed effect (see equations 14 and 15), we then specify a regression of the estimated origin FEs on local wages to obtain estimates of parameters γ and ψ_i:

$$\widehat{\mathrm{FE}}_i^o = -\frac{1}{1-\gamma}\log\left(\frac{w_i}{w_1}\right) + \frac{1}{1-\gamma}\log\left(\frac{\psi_i}{\psi_1}\right). \tag{18}$$

Since wages are not observable at such a fine geographical scale for the whole African continent, we use nighttime lights per capita as a proxy, assuming $\log\left(\frac{w_i}{w_1}\right) = \lambda \times \log\left(\frac{\texttt{light}_i}{\texttt{light}_1}\right)$, which gives:

$$\widehat{\mathrm{FE}}_i^o = -\frac{\lambda}{1-\gamma}\log\left(\frac{\texttt{light}_i}{\texttt{light}_1}\right) + \text{residual}_i. \tag{19}$$

Estimation of equation (19) proceeds via two-stage least squares (2SLS), where we use as an instrument for $\texttt{light}_i$ the (log of) average world prices of the most suitable crops:

$$\log\left(\overline{\text{Price}_i}\right) = \log\left(\sum_{t=1997}^{2021}\sum_{c=1}^{5}\alpha_i^c \times P_t^c\right),$$

where α_i^c is relative suitability of ethnic homeland i for crop c and P_t^c the worldwide real price of crop c, base 100 in 2000. For each ethnic homeland i, we consider the 5 most suitable crops. The nighttime light data is obtained from the harmonized global nighttime light dataset (Li and Zhou, 2017), population data is obtained from WorldPop[14], and crop suitability data and annual

[14]For the population data, we make use of the top-down unconstrained estimation modeling approach for 2000. See `https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained/`

real prices are obtained from FAO Agromaps and the World Bank, respectively.

Table 5 provides estimates of the instrumented regression with column (1) presenting the OLS regression, column (2) the first stage, column (3) is the reduced form version of the regression, and column (4) provides the IV estimate of (19). The coefficient is -0.875, which is $-\frac{\lambda}{1-\gamma}$ in the theory. The parameter λ is calibrated based on Bruederle and Hodler (2018) who estimate the correlation between wages and nighttime light to be 0.27. Our estimate of γ is therefore $\hat{\gamma} = 0.69$, which is consistent with negative effects of violence frictions on violence flows (a minimal result for the model not to be rejected by the data, see exponent on ζ_{in} in equation 14). The external source of violence ψ_1 is obtained from the violence-to-soldiers ratio of the Zaghawa group in Sudan-Tchad ($\psi_1 = 206/35'000 = 0.005$),[15] and the residual of the 2SLS regression can be used to recover the values of ψ_i.

Table 5: Two Stage Least Squares determinants of origin FE ($\widehat{\texttt{FE}}_i^o$)

	(1)	(2)	(3)	(4)
	OLS	First stage	Reduced form	2SLS
Dep. Var.	$\widehat{\texttt{FE}}_i^o$	$\log\left(\frac{\texttt{light}_i}{\texttt{light}_1}\right)$	$\widehat{\texttt{FE}}_i^o$	$\widehat{\texttt{FE}}_i^o$
$\log\left(\frac{\texttt{light}_i}{\texttt{light}_1}\right)$	–0.327**			–0.875*
	(0.133)			(0.452)
$\log\left(\frac{\overline{Price}_i}{\overline{Price}_1}\right)$		8.546***	–7.474**	
		(2.428)	(3.491)	
R-squared	0.055	0.127	0.050	
Observations	77	77	77	77
First-stage F-statistic				12

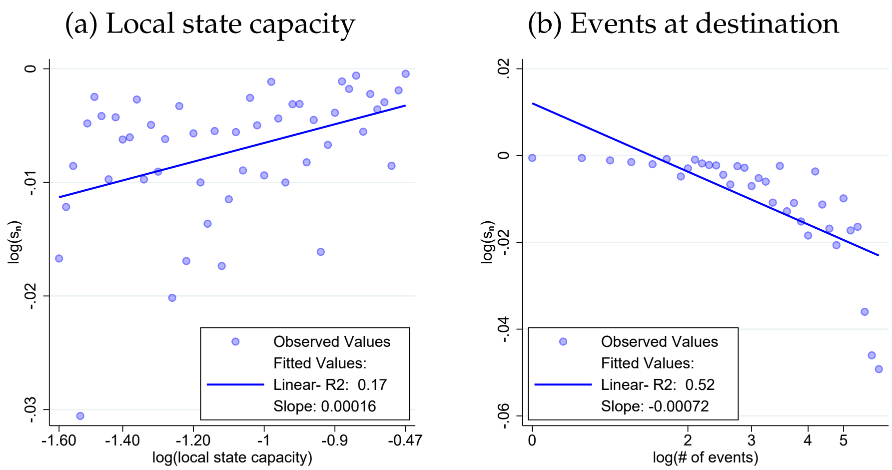
Note: Column (1) presents the OLS regression, column (2) the first stage, column (3) the reduced form version of the regression, and column (4) provides the IV estimate of Equation (19). * Significant at 10%, ** significant at 5%, *** significant at 1%. $\widehat{\texttt{FE}}_i^o$ are estimated using Equation (17). $\log\left(\frac{\texttt{light}_i}{\texttt{light}_1}\right)$ corresponds to log of the nighttime light per capital normalized by the value associated to the reference group. $\log\left(\frac{\overline{Price}_i}{\overline{Price}_1}\right)$ corresponds to the log average world prices of the most suitable crops normalized by the value of the reference group.

Step 3: The unsecured share of income can be calculated using the fact that the sum of revenues captured by i's fighters in n adds up to lost income of n:

$$\sum_i w_i l_{in} = (1 - s_n) w_n \bar{L}_n. \tag{20}$$

[15]The *Justice and Equality Movement* armed group that is affiliated to the Zaghawa group (35'000 fighters) for which we record 206 single events of violence

Figure 4: Checking $\hat{s}_n$ against state capacity and local violence



Note: Both panels present binscatter plots. Panels (a) and (b) display 50 and 35 bins respectively. Panel (a) uses data on local state capacity from Agneman et al. (2022). The estimated value of state capacity, $\hat{s}_n$, is recovered using Equation (21).

We can isolate s_n in (20), use our relationship between wages and nightlights ($w_i = \texttt{light}_i^{\lambda}$), and the definition of $\texttt{violence}_{in} \equiv \psi_i l_{in}$ to recover an estimate of state capacity as:

$$\hat{s}_n = 1 - \frac{\sum_i \left(\texttt{light}_i\right)^{\lambda} \frac{\texttt{violence}_{in}}{\widehat{\psi}_i}}{\left(\texttt{light}_n\right)^{\lambda} \texttt{pop}_n}, \tag{21}$$

where $\bar{L}_n$ is approximated by total population $\texttt{pop}_n$, ensuring that all elements of the right-hand-side of the equation are estimated, calibrated or observed.

Inspection of the distribution of $\hat{s}_n$ is an indirect validity check of our model. We can first note that all estimates of those secured shares are within the $(0, 1)$ range. Figure 4 plots $\hat{s}_n$ against state capacity and local violence. Data on state capacity comes from Agneman et al. (2022), who use i) perception of extractive, coercive, and administrative capacity (from Afrobarometer), and ii) machine learning (satellite imagery, infrastructure, geographical data...), to capture the structural impediments and incentives to state-building. Panel (a) of the figure shows a clear positive slope. In panel (b), we verify the intuitive prediction that the number of violent events should fall with the capacity of the state to ensure that a high share of income is secured.

Step 4: We recover the number of farmers/fighters in each region as:

$$\widehat{l}_i = \frac{\texttt{violence}_i}{\widehat{\psi}_i} \quad \text{and} \quad \widehat{L}_i = \texttt{pop}_i - \widehat{l}_i. \tag{22}$$

As a sanity check we can inspect the share of fighters obtained as $\widehat{l}_i / \texttt{pop}_i$. The median value is 0.04, with a standard deviation of 0.08. Furthermore we have $\widehat{L}_i > 0$ in all regions.

Step 5: Bilateral trade costs τ are calibrated using external data from Head and Mayer (2014) with a specification where trade impediments depend on distance and (national) border crossing. From the meta-analysis in Head and Mayer (2014), we set the trade elasticity $1 - \sigma = -4$. The same source provides a median elasticity of bilateral trade to distance of -1.1 and a tariff ad valorem equivalent of national borders of 60%. This enables to compute

$$\tau_{in} = \texttt{dist}_{in}^{1.1/4} \times (1 + 0.6 \times \text{border}_{in}). \tag{23}$$

Step 6: The last step is to recover farming productivity $\widehat{A_i}$. This can be done using equation (10), where we replace unobservables with recovered parameters described in the first five steps, and manipulate to write as:

$$A_i^{1-\sigma} = \sum_n \frac{\tau_{in}^{-(\sigma-1)} \left(\texttt{light}_i\right)^{\lambda\times(1-\sigma)}}{\sum_k \tau_{kn}^{-(\sigma-1)} A_k^{\sigma-1} \left(\texttt{light}_k\right)^{\lambda\times(1-\sigma)}} \widehat{s}_n \left(\frac{\texttt{light}_n}{\texttt{light}_i}\right)^{\lambda} \frac{\texttt{pop}_n}{\widehat{L}_i}. \tag{24}$$

To recover agricultural productivity $\widehat{A_i}$, we use an iterative fixed-point procedure: starting with an initial vector of A_i, which we set as the vector of nighttime lights, one can compute the right-hand side of (24), which gives a new value for the each of i's productivity, A_i'. We use a dampening factor δ to update the new vector of productivity as $\delta \times A_i + (1-\delta) \times A_i'$. Iterating those steps until the vector converges provides a fixed-point equilibrium vector of $\widehat{A_i}$.

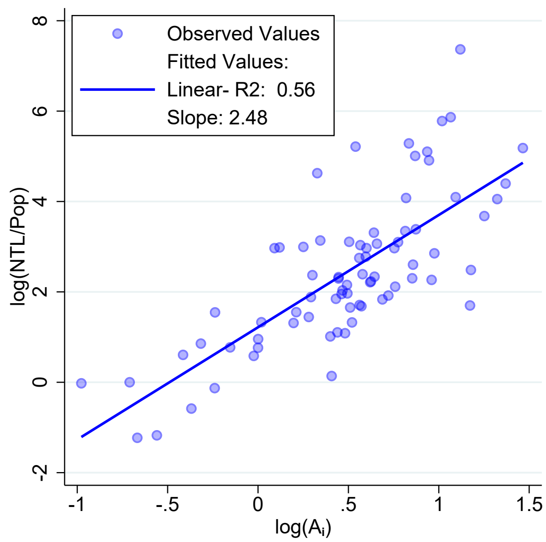
Figure 5 examines the correlation between $\widehat{A_i}$ and nighttime light. The intuition behind this check is that higher nighttime light levels (a proxy for wages in our model) should be associated with higher agricultural productivity. Indeed, we find a strong positive correlation, which confirms the internal consistency of our inversion procedure.

5 Counterfactual Simulation

5.1 Methodology

Our approach to counterfactual analysis is to shock a number of exogenous variables, most notably bilateral impediments to trade and/or violence across ethnic regions in Africa, and compute the new equilibrium to be compared with the initial one. This approach has been labeled the Difference in Expected Values (DEV) approach to counterfactual analysis by Head and Mayer (2019), since it relies on what *should be* the equilibrium in expectation according to the model evaluated at observed values of covariates and estimated parameter (hence the covariates-based approach terminology used by Dingel and Tintelnot (2020)). The alternative approach is to use the multiplicative nature of the model to obtain predicted changes of the main outcomes when imposing a policy change on *actually observed values* of those outcomes. This is called the Exact Hat Algebra (EHA) approach in the literature since Dekle et al. (2007) (the "calibrated-shares procedure" for Dingel and Tintelnot (2020)). While the debate is still on regarding the respective virtues of the two approaches, we do not have the luxury of a choice here. Since EHA relies on observed market shares, we cannot implement it in our context, since we do not observe shares of trade flows (π_{in})

Figure 5: Correlation of estimated productivity with (log-) night lights



Note: The log of the estimated farming productivity $\widehat{A_i}$ on the x-axis is obtained from equation (24).

between the regions of interest in Africa. We therefore have to first compute the expected equilibrium (which differs from the one actually observed) and compare it to the expected outcome after the policy shock

Section 4 provides the steps to estimate and "invert" the model, such that we can recover all parameters needed to compute the expected equilibrium. The set of structural parameters (observed, calibrated or recovered from model estimation/inversion) used to compute equilibrium is given by:

$$\Theta = \{\gamma, \sigma, s_n, \psi_i, \text{pop}_i, A_i, \tau_{in}, \xi_{in}\}.$$

Equilibrium at observed values of Θ is obtained by a fixed point iteration on equation (25) which transforms an initial vector of wages $\mathbf{w}$ into a final vector $\mathbf{w'}$:

$$w'_i = \sum_n \beta_{in}(\mathbf{w}) w_n \frac{\text{pop}_n}{\text{pop}_i}, \tag{25}$$

where $\beta_{in}(.)$ is a non-linear function of the wage vector $\mathbf{w}$ given by equation (13). The algorithm used to find the initial equilibrium is then

Wage (inner) loop steps:

0: Use $w_i = \frac{A_i}{A_1}$ as values for initializing the procedure.

1: With the vector of w_i (normalized such that $w_1 = 1$) obtained from the previous iteration, compute the RHS of Equation (25). The LHS of (25) yields a new vector w'_i (also to be normalized by w'_1).

2: Use a dampening factor δ to compute the new vector of wages as $\delta \times w_i + (1-\delta) \times w_i'$.

Steps 1 and 2 are repeated until the vector of wages stops changing (up to a level of tolerance).

The fixed point iteration on (25) is sufficient for computing the initial equilibrium (with "observed" productivity provided by equation 24). This is because the equilibrium vector of wages conditional on Θ is sufficient to compute flows, allocations of individuals into farming vs fighting and of fighters over destinations and welfare:

$$w_n/P_n = w_n / \left[\sum_k \left(\frac{\tau_{kn} w_k}{A_k}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{26}$$

A counterfactual change (in τ for instance) simply re-runs the algorithm on (25) and obtains new outcomes $\texttt{violence}'_{in}$, w_n' and P_n' in particular that can be compared to $\texttt{violence}_{in}$, w_n and P_n. However, this approach to counterfactual analysis only works for a setup where productivity in n, A_n, is not affected by violence taking place there, and to the changes in violence that the model predicts under the new equilibrium. Given the wealth of evidence regarding the large impact of conflictuality on output and income, it seems unrealistic to hold the vector $\mathbf{A}$ constant in our simulations. In the counterfactual simulations presented below, the model is therefore augmented with "destruction spillovers", such that

$$A_n = \bar{A}_n \exp\{-\varepsilon \times \texttt{violence}_n\,[\mathbf{w}(\mathbf{A})]\}, \tag{27}$$

where the parameter ε denotes the semi-elasticity of TFP with respect to violence.[16] Note that as soon as productivity depends on violence, solving for the vector of wages is not sufficient anymore to compute equilibrium, since wages are both affected by productivities (equation 13), and feedback to impact them through equilibrium violence (equation 27). We therefore use an inner-outer loop fixed point procedure in our counterfactual scenarios. The inner loop characterizes the equilibrium wage vector $\mathbf{w}$, given a productivity vector $\mathbf{A}$ (as described above), while the outer loop takes the resulting $\mathbf{w}$ vector from the inner loop to find the equilibrium productivity vector $\mathbf{A}$.

Productivity (outer) loop steps:

1: Use the vector of w_i obtained from the inner loop, compute the RHS of Equation (27). The LHS of (27) yields a new vector A_n'.

2: Use a dampening factor δ to compute the new vector of productivities as $\delta \times A_n + (1-\delta) \times A_n'$.

[16] Calibration of ε is based on a 2SLS regression of the log of A_n on $\texttt{violence}_n$, with violence being instrumented by a model-driven supply shifter of violence $\sum_i \left(\frac{\xi_{in}}{\psi_i}\right)^{-\frac{\gamma}{1-\gamma}}$. The calibrated semi-elasticity of TFP to violence is found to be $\varepsilon = 0.0075$ (one additional ACLED event reduces the local TFP by 0.75%). To recover the baseline productivity level $\bar{\mathrm{A}}_n$, the calibrated ε, estimated A_n, and observed $\texttt{violence}_n$ are used to invert equation (27).

The overall simulation procedure stops when the tolerance level for changes in the vector of $\mathbf{A}$ is reached.

The welfare impact of the counterfactual scenarios that we simulate can be related to a sufficient statistics formula well-known in trade since the work of Arkolakis et al. (2012). We start by writing the level of welfare before the policy change as $\omega_n = \frac{s_n w_n}{P_n}$. The share of trade with self in region n is given by $\pi_{nn} = (\tau_{nn} w_n / A_n)^{1-\sigma} P_n^{\sigma-1}$ meaning that $\frac{w_n}{P_n} = \pi_{nn}^{\frac{1}{1-\sigma}} \frac{A_n}{\tau_{nn}}$. Assuming that τ_{nn}, s_n and $\bar{A}_n$ are unaffected by the counterfactual experiment:

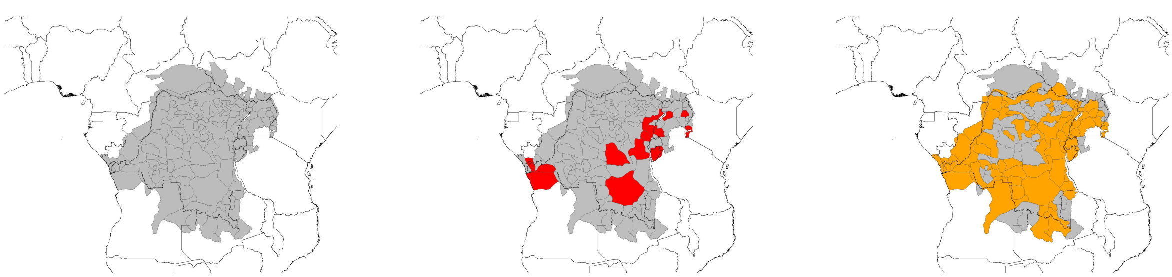
$$\frac{\omega_n'}{\omega_n} = \left(\frac{\pi_{nn}'}{\pi_{nn}}\right)^{\frac{1}{1-\sigma}} \frac{A_n'}{A_n} = \left(\frac{\pi_{nn}'}{\pi_{nn}}\right)^{\frac{1}{1-\sigma}} \exp\{-\varepsilon(\texttt{violence}_n' - \texttt{violence}_n)\}.$$

The first term is the traditional formula of Arkolakis et al. (2012), which here includes an indirect influence of violence on welfare through changes in the trade patterns. The second term is an additional direct effect that accounts for the equilibrium change of violence and therefore induced increase or decrease in the level of damage to productivity.

5.2 World Bank Great Lakes Initiative (GLI)

Our first natural counterfactual experiment is inspired by the World Bank's Great Lakes Initiative (GLI), which has a dual objective of achieving development and peace in the Great Lakes Region (GLR). The area comprises 131 ethnic regions spanning over Democratic Republic of Congo (DRC), Rwanda, and Uganda, out of a total of 824 regions in Africa. Violence within GLR has represented 32% of all violence in Africa between 1997 and 2022, with 14 attacking ethnic regions and 69 attacked ethnic regions (figure 6 provides a map of those events). The GLI seeks to address the root causes of conflict in the region and promote sustainable development, with a focus on improving governance, promoting economic growth, and increasing access to basic services such as health and education. The initiative also supports cross-border collaboration and regional integration to strengthen economic interdependence and promote stability.[17]

Figure 6: Violence in the Great Lakes Region over our period



131 Ethnic groups in GLR | 14 Origins of violence | 69 Destinations of violence

Facilitation of cross-border trade through infrastructure and capacity building is a key component of the GLI.[18] Those infrastructures and simpler cross-border passage are however also sus-

[17] https://ungreatlakes.unmissions.org/sites/default/files/world_bank_great_lakes_initiative.pdf

[18] As a part of the wider GLI, the Great Lakes Trade Facilitation Project (GLTFP) (P151083) was approved by the Board

ceptible of reducing frictions to violence. We therefore build a counterfactual scenario cutting both frictions. Our benchmark takes seriously the ambition of the plan by halving the Ad Valorem Equivalents (AVE) of RDC-RWA-UGA borders crossing for trade frictions and fighting frictions (initially at 60% and 87% respectively, see section 4):

$$\tau_{in} = \texttt{dist}_{in}^{1.1/4} \times (1 + 0.6 \times \text{border}_{in}) \quad \rightarrow \quad \tau'_{in} = \texttt{dist}_{in}^{1.1/4} \times (1 + 0.3 \times \text{border}_{in}) \tag{28}$$

$$\xi_{in} = \text{dist}_{in}^{\alpha_1} \times (1 + 0.87 \times \text{border}_{in}) \quad \rightarrow \quad \xi'_{in} = \text{dist}_{in}^{\alpha_1} \times (1 + 0.435 \times \text{border}_{in}) \tag{29}$$

The " intervention" on τ leads to a reduction in the trade-reducing effect of national borders which goes from -85% ($= 1.6^{1-\sigma} - 1$) to -65% ($= 1.3^{1-\sigma} - 1$). The corresponding fall in frictions to violence leads the negative impact of border-crossing on violence to go from -75% ($= 1.87^{-\gamma/(1-\gamma)} - 1$) to -55% ($= 1.435^{-\gamma/(1-\gamma)} - 1$).

Table 6 reports a number of important statistics predicted by the model in the initial (pre-GLI) equilibrium. The two rows average across ethnic regions in the Great Lakes Region or Rest Of Africa, considered as destinations n = GLR and n = ROA respectively. The two first columns show the origin of violence as predicted by the model: while RoA attacks are not unfrequent, most of the events that occur in the GLR originate from inside the region. The next four columns compute the expected degree of trade integration between our spatial units of analysis. Slightly more that 15% of trade takes place within an ethnic region ($i = n$) in our model. Nearly 23% of expenditure is spent on goods from a group that belongs to another country within the GLR region (i and n both in GLR but not in the same country). Those pairs are the target of the policy experiment, intended to reduce the cost faced by trade flows when crossing a national border. Overall, other GLR origins (with $i \neq n$) represent more than 60% of GLR ethnic groups purchases outside of self-trade, and Rest of Africa less than a quarter of those purchases. Note that those trade patterns are entirely driven by the structure and parameters imposed on the model, since trade flows are not observed at that level of detail.

Table 6: Initial equilibrium: some statistics

	Avg. % violence		Average percent imports			
Origin:	GLR	RoA	All Groups	Foreign Groups	GLR	RoA
Destination:						
GLR	59.03	40.97	86.49	8.15	71.42	15.07
RoA	1.93	98.07	84.63	8.64	4.94	79.69

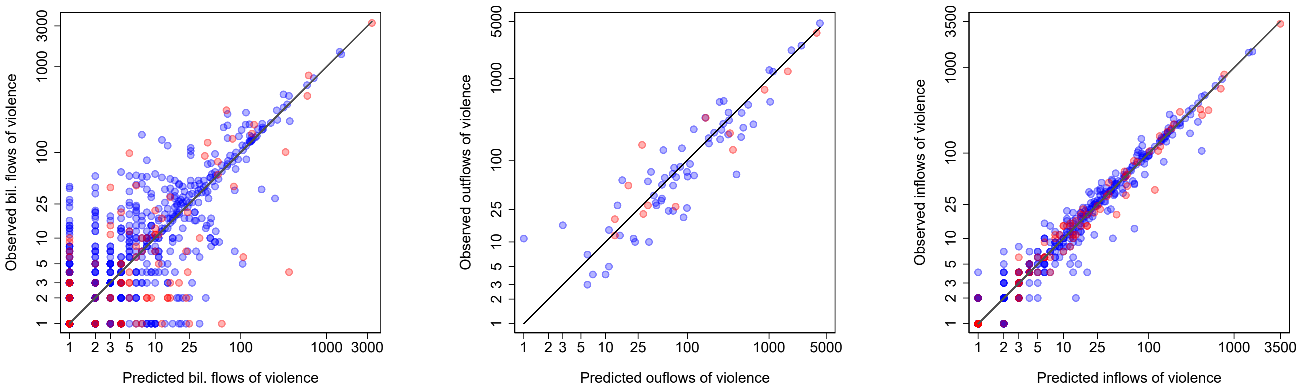
Note: Numbers represent average percentage points. GLR = Great Lakes Region, and RoA = Rest of Africa.

For bilateral violence however, we can confront the levels predicted by the model with the

of the World Bank on September 25, 2015. The project implementation stalled mostly due to budget limitation and the Covid-19 pandemic. In late 2020, none of the planned improvements to border infrastructures were realized in all three countries. Only in one site (Goma), temporary facilities had been in place since 2019. World Bank (2020) provides a detailed report on the project status and progress.

observed flows. Figure 7 provides such comparison for the whole cross-section (not restricted to GLR destinations). The three panels show in red the set of GLR countries, and in blue the rest of Africa. Panel (a) reports on the x-axis bilateral flows of violence between i and n, while the y-axis reports the observed corresponding flow (in log scales and rounded to the nearest unit in both cases). Although deviations are larger for "small" cells where idiosyncratic shocks to violence matter more as should be expected, the model does an overall good job at predicting the spatial pattern of violence. The total violence generated by i (panel b) or suffered by n (panel c)are even better predicted.

Figure 7: Goodness of fit



(a) Bilateral (b) Outward (c) Inward

Table 7 provides a first set of results for three different simulations corresponding to three versions of the model (reported in the first column). The first version ("Pure Trade") neutralizes the changes in violence that would otherwise be happening following the counterfactual policy change. This is done by imposing $l'_{in} = l_{in}$ (see the appendix for details). As a consequence, this scenario is a useful comparison point with a model that would not consider endogenous response of violence to changes in economic incentives following a trade liberalization episode. The usual Arkolakis et al. (2012) formula applies, in which the change in welfare is a function of the change in domestic flows raised to the inverse of the trade elasticity. Domestic flows fall by an average near nine percent following the trade liberalization with neighbors inside the GLR, and welfare on average increases by 2.3 percent.

The second incarnation of the model ("Damage-free") allows violence flows to react to the reduced effects of national borders but freezes the feedback loop to TFP, by setting $\varepsilon = 0$ in equation (27), as shown in the tenth column of the table where the average change in TFP remains a uniform 0. A first important finding is that the overall violence in the region hardly changes, which is shown in the third column of the table as a small 0.1% increase in total violence in GLR. As a consequence, overall trade patterns and welfare are almost unchanged. Columns 4 to 6 give changes in bilateral violence for the average in pair, again distinguishing over whether origins and destinations belong to GLR or RoA. The average GLR to GLR violence flows goes up by 16%. There are many mechanisms behind this number. In order to quantify those mechanisms, it is useful to come back to the

Table 7: Counterfactual equilibrium: results

% Change in :		Agg Violence	Avg violence			Avg trade share			Average	
			All	GLR	RoA	Self	GLR	RoA	TFP	Welfare
Model	Region									
Pure Trade	GLR	0	0	0	0	-8.55	9.22	-4.64	0	2.29
	RoA	0	0	0	0	.35	-4.25	.32	0	-.09
Damage-free	GLR	.1	1.14	16.05	-8.33	-8.55	9.22	-4.63	0	2.29
	RoA	-.13	-.08	-3.88	0	.35	-4.26	.32	0	-.09
Damage-inclusive	GLR	-1.01	-1.06	8.56	-3.88	-8.76	9.24	-4.88	.4	2.76
	RoA	-.09	-.21	-16.97	.2	.29	-3.69	.29	.01	-.06

Note: Numbers represent average percent changes. GLR = Great Lakes Region, and RoA = Rest of Africa

gravity of violence equation (6), and consider the determinants behind a change in a bilateral flow.

$$\Delta \log \texttt{violence}_{in} = -\frac{\gamma}{1-\gamma} \Delta \log \xi_{in} - \frac{1}{1-\gamma} \Delta \log w_i + \Delta \log w_n - \Delta \log \mathrm{MRV}_n. \qquad (30)$$

The first term in equation (30) is the most obvious: with a decrease in frictions of violence (involuntarily) associated with the GLI, bilateral violence increases between groups that belong to the great lakes region but are in different countries. The two next terms are wage terms. When the policy increases nominal wages at origin, it reduces the flow of violence, because it drives potential fighters to stay in (or move to) farming. Conversely, an increase of income at destination increases the incentives of all fighters to concentrate their rapacity efforts on n. Note that in the case of internal violence, $w_i = w_n$ and the two wage terms become $-\frac{\gamma}{1-\gamma} \Delta \log w_n$. Because our estimates give $0 < \gamma < 1$, an increase in wages reduces internal violence more than proportionately. The last term is the change in the multilateral resistance of violence. It captures the intensity of competition on destination of violence n.

Table 8: Damage-free counterfactual equilibrium: decomposition

i and n are		#	Base	Average Δ log				
in diff. countries	$\neq$	dyads	$\texttt{violence}_{in}$	$\texttt{violence}_{in}$	$\xi_{in}^{-\frac{\gamma}{1-\gamma}}$	$w_i^{-\frac{1}{1-\gamma}}$	w_n	MRV_n^{-1}
0	0	13	336.9	-.097	0	-.037	.011	-.071
0	1	414	5.1	-.101	0	-.033	.009	-.076
1	1	777	.9	.448	.592	-.037	.013	-.12

Note: Numbers represent average percentage point changes for the ethnic groups considered as destinations of either violence or trade (n). GLR = Great Lakes Region, and RoA = Rest of Africa

Table 8 provides the decomposition for the changes in bilateral flows of violence that occur within the GLR region in the Damage-free scenario (the 16% increase in row 3 and column 5 of

Table 7). The first row is the simplest: it decomposes the change in local violence inside the 13 ethnic groups that fight internally. While this is a very small share of the dyads, their initial level of violence is extremely high. For those groups, change in violence stems from two simple effects: the increase in wage reduces violence (as explained above), and the increase in competition from neighboring regions that face a fall in their cost of raiding n also reduces the incentives of fighters in n to loot locally. The mechanisms are very similar in the second row of Table 8, which is the case of two different groups in the same country. The third row of the table shows the quantitative importance of violence frictions since those are dyads separated by a national border. The increase in violence generated by this determinant dominates largely the other ones (which remain collectively violence-reducing). Even though the initial count of events is very low in those pairs, their large increase of violence brings the overall increase of violence to 16%.

Table 9: Damage-inclusive counterfactual equilibrium: decomposition

i and n are		#	Base	Average Δ log				
in diff. countries	$\neq$	dyads	$\texttt{violence}_{in}$	$\texttt{violence}_{in}$	$\xi_{in}^{-\frac{\gamma}{1-\gamma}}$	$w_i^{-\frac{1}{1-\gamma}}$	w_n	MRV_n^{-1}
0	0	13	336.9	-.074	0	-.1	.031	-.005
0	1	414	5.1	-.08	0	-.059	.011	-.032
1	1	777	.9	.415	.592	-.113	.019	-.082

Note: Numbers represent average percentage point changes. GLR = Great Lakes Region, and RoA = Rest of Africa

Note that the quantitative importance of frictions just mentioned explains the *fall in violence originating from RoA* (-8.33% in Table 7). Indeed when $i \in$ RoA, fighters do not "benefit" from reduced frictions to attack GLR groups and only face increased competition (increased MRV_n) when trying to raid regions affected by the GLI. This dominates the increased rapacity incentives coming from the fact that groups in GLR are made richer by the policy experiment.

In the last version of the model ("Damage-inclusive" in Table 7), the regions of the GLR are made even more well-off by the policy since TFP increases by 0.4% (which raises welfare gains close to 2.8%). Hence the rapacity incentives is stronger making the average fall in violence for this pair more modest (-3.88%). Table 9 provides the same decomposition as Table 8, but account for damage spillovers. The main difference between the two tables lies within the third row. The wage-at-origin effect is tripled for groups facing a reduction in violence frictions at the national border, which reduces the overall violence increase for those pairs. This is enough to turn the overall violence effect of the GLI negative, which reinforces the welfare gains of GLI: the trade and violence channels turn out to be complement in the full version of our model.

References

Agneman, G., K. Brandt, C. Cappelen, and D. Sjöberg (2022). Mapping local state capacity. SSRN 4184751.

Allen, T., C. Arkolakis, and Y. Takahashi (2020). Universal gravity. *Journal of Political Economy 128*(2), 393 – 433.

Alvarez, F. and R. E. Lucas (2007). General equilibrium analysis of the eaton–kortum model of international trade. *Journal of monetary Economics 54*(6), 1726–1768.

Amarasinghe, A., P. Raschky, Y. Zenou, and J. Zhou (2020). "conflicts in spatial networks. *mimeo Monash University*.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review 93*(1), 170–192.

Anderson, S., A. De Palma, and J. Thisse (1992). *Discrete choice theory of product differentiation*. MIT Press.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers - International Monetary Fund 16*(1), 159–178.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2017, June). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review 107*(6), 1564–1610.

Berman, N., M. Couttenier, and R. Soubeyran (2021). Fertile Ground for Conflict. *Journal of the European Economic Association 19*(1), 82–127.

Bruederle, A. and R. Hodler (2018, 09). Nighttime lights as a proxy for human development at the local level. *PLOS ONE 13*(9), 1–22.

Buhaug, H., K. S. Gleditsch, H. Holtermann, G. Østby, and A. F. Tollefsen (2011). It's the local economy, stupid! geographic wealth dispersion and conflict outbreak location. *Journal of Conflict Resolution 55*(5), 814–840.

Buhaug, H. and J. K. Rød (2006). Local determinants of african civil wars, 1970–2001. *Political Geography 25*, 315–335.

Burke, M., S. M. Hsiang, and E. Miguel (2015). Climate and conflict. *Annual Review of Economics 7*(1), 577–617.

Cederman, L.-E., H. Buhaug, and J. K. RÃ¸d (2009, August). Ethno-Nationalist Dyads and Civil War. *Journal of Conflict Resolution 53*(4), 496–525.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–21.

Collier, P. and A. Hoeffler (2004, October). Greed and grievance in civil war. *Oxford Economic Papers 56*(4), 563–595.

Combes, P.-P., M. Lafourcade, and T. Mayer (2005). The trade-creating effects of business and social networks: evidence from france. *Journal of International Economics 66*(1), 1–29.

Crost, B., J. Felter, and P. Johnston (2014, June). Aid under fire: Development projects and civil conflict. *American Economic Review 104*(6), 1833–1856.

Dal Bó, E. and P. Dal Bó (2011, August). Workers, Warriors, And Criminals: Social Conflict In General Equilibrium. *Journal of the European Economic Association 9*(4), 646–677.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced Trade. *American Economic Review 97*(2), 351–355.

Dell, M., B. F. Jones, and B. A. Olken (2014, September). What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature 52*(3), 740–98.

Dingel, J. and F. Tintelnot (2020). Spatial economics for granular settings. Technical report, National Bureau of Economic Research.

Dube, O. and J. Vargas (2013). Commodity price shocks and civil conflict: Evidence from colombia. *Review of Economic Studies 80*(4), 1384–1421.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eberle, U. J., D. Rohner, and M. Thoenig (2020, December). Heat and Hate: Climate Security and Farmer-Herder Conflicts in Africa. CEPR Discussion Papers 15542, C.E.P.R. Discussion Papers.

Fally, T. (2012, June). Structural gravity and fixed effects. University of Colorado note.

Harari, M. and E. L. Ferrara (2018, October). Conflict, Climate, and Cells: A Disaggregated Analysis. *The Review of Economics and Statistics 100*(4), 594–608.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse,toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of international economics*, Volume 4, pp. 131–196. Elsevier.

Head, K. and T. Mayer (2019). Brands in Motion: How frictions shape multinational production. *American Economic Review 109*(9), 3073–3124.

König, M. D., D. Rohner, M. Thoenig, and F. Zilibotti (2017). Networks in conflict: Theory and evidence from the great war of africa. *Econometrica 85*(4), 1093–1132.

Korovkin, V. and A. Makarin (2021). Production networks and war: Evidence from ukraine. CEPR Discussion Papers 16759, C.E.P.R. Discussion Papers.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review 70*(5), 950–959.

Li, X. and Y. Zhou (2017). A stepwise calibration of global dmsp/ols stable nighttime light data (1992–2013). *Remote Sensing 9*(6).

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. New York: Oxford University Press.

McGuirk, E. and M. Burke (2020a). The economic origins of conflict in africa. *Journal of Political Economy 128*(10), 3940–3997.

McGuirk, E. and M. Burke (2020b). The Economic Origins of Conflict in Africa. *Journal of Political Economy 128*(10), 3940–3997.

McGuirk, E. F. and N. Nunn (2020, December). Transhumant Pastoralism, Climate Change, and Conflict in Africa. NBER Working Papers 28243, National Bureau of Economic Research, Inc.

Moscona, J., N. Nunn, and J. A. Robinson (2020). Segmentary lineage organization and conflict in sub-saharan africa. *Econometrica 88*(5), 1999–2036.

Mueller, H., D. Rohner, and D. Schönholzer (2022). Ethnic Violence Across Space ['A theoretical foundation for the gravity equation']. *The Economic Journal 132*(642), 709–740.

Murdock, G. P. (1959). An Atlas of African History. JD Fage. *American Anthropologist 61*(3), 530–531.

Nunn, N. and N. Qian (2014, June). Us food aid and civil conflict. *American Economic Review 104*(6), 1630–66.

OECD (2009). *Armed violence reduction: Enabling development*. OECD.

Olson, M. (1993). Dictatorship, democracy, and development. *American Political Science Review 87*(3), 567–576.

Proost, S. and J.-F. Thisse (2019, September). What can be learned from spatial economics? *Journal of Economic Literature 57*(3), 575–643.

Redding, S. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1), 21–58.

Sanchez de la Sierra, R. (2020). On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo. *Journal of Political Economy 128*(1), 32–74.

Tilly, C. (1985). *War Making and State Making as Organized Crime*, pp. 169–191. Cambridge University Press.

World Bank (2020). *Proposed project restructuring of AFR Great Lakes Trade Facilitation (P151083), Report no: RES38891*. World Bank.

World Bank Group (2020). Commodity markets outlook, october 2020. https://openknowledge.worldbank.org/handle/10986/34621 , License: CC BY 3.0 IGO.

Appendix

A The appropriation game

Setup. Appropriation is modeled as a multi-stage game with repeated contests. The sequence works as follows: farmers are first looted by fighters who grab $(1-s_n)Y_n^P$. Then, fighters are themselves looted by other fighters who are looted by other fighters... until the game ends. The game ends when income is ε-secured (i.e. unsecured income falls below an arbitrary $\varepsilon > 0$) for *all* players. The box below describes the game:

<u>Sequence of the appropriation game:</u>

1. Front-loaded payment of the gross income Y_n^P of producers; only a share s_n is immediately (and definitively) secured. The residual $(1-s_n)Y_n^P$ is unsecured.
2. "Once-for-all" optimal assignment of fighters l_{in} from region i to region $n \rightarrow$ this sets the stationary appropriation shares p_{in} as defined by the CSF and subject to a spatial friction factor ξ_{in}.
3. <u>Sub-period 1:</u> Fighters in i loot unsecured farmers' income, generating revenue $R_i(1) = \sum_n p_{in} \times (1-s_n)Y_n^P$.
4. Repeated stage game starts:

   <u>Stage Game at sub-period $k > 1$:</u>

   i/ Looting by fighters l_{in} of income that is still unsecured in region n.

   ii/ $R_i(k)$ is the flow of income appropriated by fighters in sub-period k. It is (friction-free) repatriated in region i.

   iii/ A share s_i of $R_i(k)$ is *definitively* secured. The residual income $(1-s_i)R_i(k)$ is unsecured.

   iv/ If $(1-s_i)R_i(k) < \varepsilon$ for all i, the sub-game ends and we move to stage 5 (below).[19] Otherwise, we proceed to sub-period $(k+1)$ and restart the stage game (i) to (iii).
5. Secured incomes of all resident workers from i (farmers and fighters) being repatriated in i, production, trade, and consumption take place.

Accounting exercise: "Follow the money". To understand how looting of resources evolves over time, we need to perform an accounting of the looted resources at each stage of the game. The law of motion of appropriation is given by

$$R_i(1) = \sum_n p_{in} \times (1-s_n)Y_n^P, \quad \text{and} \quad R_i(k) = \sum_n p_{in} \times (1-s_n)R_n(k-1), \quad \text{for} \quad k > 1. \tag{A.1}$$

In matrix notation:

$$\mathbf{R}(1) = \mathbf{A}\mathbf{Y}^{\mathbf{P}}, \quad \text{and} \quad \mathbf{R}(k) = \mathbf{A}\mathbf{R}(k-1) \text{ for } \quad k > 1, \tag{A.2}$$

where $\mathbf{A}$ is the $(N \times N)$ appropriation matrix: $a_{in} = p_{in} \times (1-s_n)$.

Replacing $\mathbf{R}(k-1) = \mathbf{AR}(k-2)$, $\mathbf{R}(k-2) = \mathbf{AR}(k-3)$, etc. in equation (A.2) yields

$$\mathbf{R}(k) = \mathbf{A}^k \mathbf{Y}^\mathbf{P}. \tag{A.3}$$

As k grows large, the vector $\mathbf{R}(k)$ converges to the null vector (this is because all entries of A are positive and below 1).

The amount of *gross fighting revenues* accumulated over the entire game is given by

$$\mathbf{R} = \sum_{k=1}^{\infty} \mathbf{R}(k) = \sum_{k=1}^{\infty} \mathbf{A}^k \mathbf{Y}^\mathbf{P} = \mathbf{A}\left(\mathbf{Y}^\mathbf{P} + \sum_{k=1}^{\infty} \mathbf{A}^k \mathbf{Y}^\mathbf{P}\right) = \mathbf{A}\left(\mathbf{Y}^\mathbf{P} + \mathbf{R}\right) = \mathbf{AY}, \tag{A.4}$$

where $\mathbf{Y} \equiv \mathbf{Y}^\mathbf{P} + \mathbf{R}$ is the vector of total gross incomes (farmers' + fighters' incomes). From the previous equation, we obtain the gross fighting revenues accruing to region i

$$R_i = \sum_n p_{in} \times (1 - s_n) Y_n \tag{A.5}$$

This revenue is the one that each group seeks to optimize by assigning optimally its fighters l_{in} to each region n.

From gross income to expenditure. An important distinction in the model is between aggregate gross income Y_n that is made of *both* secured and unsecured income and aggregate expenditure E_n that is made *only* of fully secured income.

Gross aggregate income is given by producers' and farmers' incomes

$$Y_i = Y_i^P + R_i$$

Total expenditures of producers are given by

$$E_i^P = s_i Y_i^P$$

For fighters, at each subperiod k of the game, only a share of their newly appropriated flow of revenues is secured and will ultimately contribute to their expenditures. Thus, their expenditure is given by

$$E_i^F = \sum_{k=1}^{\infty} s_i R_i(k) = s_i \sum_{k=1}^{\infty} R_i(k) = s_i R_i$$

We consequently get that the total expenditure of region i is given by

$$E_i = E_i^P + E_i^F = s_i \left(Y_i^P + R_i\right) = s_i Y_i \tag{A.6}$$

Consequences for the GE. Under these micro-foundations of the appropriation game, we see from the previous relation that the aggregate ratio E_n/Y_n is equal to the exogenous parameter s_n. Essentially, the assumption that "being a fighter does not protect against looting" implies that this ratio is not affected by the (endogenous) composition of the workforce at equilibrium: it does not depend on the number of fighters and farmers.

Importantly, the relation $E_n = s_n Y_n$ proves to be **very useful** (but not vital) for the GE analysis: it leads to the symmetry in the aggregate trade and fighting revenue equations. This in turn enables us to characterize the GE with the unique fixed point "master" equation (see the main text).

General Equilibrium under alternative microfoundations. Let assume alternatively that only farmers are looted: Hence, fighters' income is immediately and fully secured. The rest of the game is unchanged. Clearly, after one step, farmers are looted and the game ends. Moreover, at equilibrium of the labor market, workers are indifferent between farming and fighting. Therefore:

$$s_i \times w_i^P = w_i^F$$

In turn, denoting w_i the fighters' wage at equilibrium, aggregate expenditures are given by

$$E_n = s_n \frac{w_n}{s_n} L_n + w_n l_n = w_n \bar{L}_n$$

And the system of equations that characterizes the GE is given by

$$\frac{w_i}{s_i} L_i = \sum_n \frac{\tau_{in}^{-(\sigma-1)} \left(\frac{s_i A_i}{w_i}\right)^{\sigma-1}}{\sum_k \tau_{kn}^{-(\sigma-1)} \left(\frac{s_k A_k}{w_k}\right)^{\sigma-1}} w_n \bar{L}_n \tag{A.7}$$

$$w_i l_i = \sum_n \frac{\xi_{in}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_i}{w_i}\right)^{\frac{\gamma}{1-\gamma}}}{\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_k}{w_k}\right)^{\frac{\gamma}{1-\gamma}}} \frac{1-s_n}{s_n} w_n L_n \tag{A.8}$$

$$\bar{L}_i = L_i + l_i \tag{A.9}$$

The system of equations (A.7)-(A.8)-(A.9) is not reducible to a unique fixed-point "master" equation anymore. So, the GE model becomes less tractable: in particular, characterizing the existence and uniqueness of the equilibrium is more challenging. This said, numerical simulations are still feasible and the counterfactual analysis could presumably be performed under this alternative model.

B Derivation of the gravity of violence

Proposition 1. *The partial equilibrium flow of violence (quantity) from i to n:*

$$\texttt{violence}_{in} \equiv \psi_i l_{in} = \xi_{in}^{-\frac{\gamma}{1-\gamma}} \times \left(\frac{\psi_i}{w_i^F}\right)^{\frac{1}{1-\gamma}} \times \frac{(1-s_n)Y_n}{\sum_k \xi_{kn}^{-\frac{\gamma}{1-\gamma}} \left(\frac{\psi_k}{w_k^F}\right)^{\frac{\gamma}{1-\gamma}}}. \tag{B.10}$$

Proof. The proof proceeds in several steps

1. The optimal spatial allocation choice of troops across regions of destination n is given by:

$$R_i \equiv \max_{l_{in}} \sum_n p_{in} \left(1-s_n\right) Y_n, \quad \text{s.t} \quad l_i = \sum_n l_{in}, \tag{B.11}$$

   with

$$p_{in} = \frac{(\xi_{in}^{-1} \psi_i l_{in})^\gamma}{\sum_k (\xi_{kn}^{-1} \psi_k l_{kn})^\gamma}, \tag{B.12}$$

   In order to obtain a concave maximization problem of R_i with respect to all optimal allocations of fighters l_{in} by group i, we need $0 < \gamma < 1$ (which we estimate to be the case in empirics). This will also ensure that an interior solution exists. The resulting Lagrangian for this optimization problem writes as:

$$\mathcal{L} = \sum_n \frac{(\xi_{in}^{-1}\psi_i l_{in})^\gamma}{\sum_k (\xi_{kn}^{-1}\psi_k l_{kn})^\gamma}(1-s_n)Y_n - \lambda\left(l_i - \sum_n l_{in}\right).$$

Assuming that groups are small enough to neglect their impact on the overall conditions of violence in n, $\frac{\partial\left[\sum_k(\xi_{kn}^{-1}\psi_k l_{kn})^\gamma\right]}{\partial l_{in}} = 0$, we obtain the first-order-conditions

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial l_{in}} = 0 &\iff l_{in} = \left(\lambda \times \frac{\sum_k \xi_{kn}^{-\gamma}\psi_k^\gamma l_{kn}^\gamma}{\gamma\xi_{in}^{-\gamma}\psi_i^\gamma(1-s_n)Y_n}\right)^{\frac{1}{\gamma-1}} \\
\frac{\partial\mathcal{L}}{\partial\lambda} = 0 &\iff l_i = \sum_n l_{in}.
\end{aligned}$$

Combining the two FOCs and using $l_i = \sum_n l_{in} = \sum_k l_{ik}$:

$$l_i = \sum_k l_{ik} \iff l_i = \lambda^{\frac{1}{\gamma-1}} \times \sum_k \left(\frac{\sum_j \xi_{jk}^{-\gamma}\psi_j^\gamma l_{jk}^\gamma}{\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma(1-s_k)Y_k}\right)^{\frac{1}{\gamma-1}} \iff \lambda^{\frac{1}{\gamma-1}} = \frac{l_i}{\sum_k\left(\frac{\sum_j \xi_{jk}^{-\gamma}\psi_j^\gamma l_{jk}^\gamma}{\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma(1-s_k)Y_k}\right)^{\frac{1}{\gamma-1}}}.$$

Plugging back this expression in the first FOC, we obtain:

$$\frac{l_{in}}{l_i} = \left(\frac{\sum_k \xi_{kn}^{-\gamma}\psi_k^\gamma l_{kn}^\gamma}{\gamma\xi_{in}^{-\gamma}\psi_i^\gamma(1-s_n)Y_n}\right)^{\frac{1}{\gamma-1}} \times \left[\sum_k\left(\frac{\sum_j \xi_{jk}^{-\gamma}\psi_j^\gamma l_{jk}^\gamma}{\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma(1-s_k)Y_k}\right)^{\frac{1}{\gamma-1}}\right]^{-1}. \tag{B.13}$$

One can further note that:

$$\left(\frac{\sum_k \xi_{kn}^{-\gamma}\psi_k^\gamma l_{kn}^\gamma}{\gamma\xi_{in}^{-\gamma}\psi_i^\gamma(1-s_n)Y_n}\right)^{\frac{1}{\gamma-1}} = (\gamma\xi_{in}^{-\gamma}\psi_i^\gamma)^{\frac{1}{1-\gamma}} \times \left(\frac{(1-s_n)Y_n}{\sum_k \xi_{kn}^{-\gamma}\psi_k^\gamma l_{kn}^\gamma}\right)^{\frac{1}{1-\gamma}},$$

and :

$$\sum_k\left(\frac{\sum_j \xi_{jk}^{-\gamma}\psi_j^\gamma l_{jk}^\gamma}{\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma(1-s_k)Y_k}\right)^{\frac{1}{\gamma-1}} = \sum_k\left[\left(\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma\right)^{\frac{1}{1-\gamma}} \times \left(\frac{(1-s_k)Y_k}{\sum_j \xi_{jk}^{-\gamma}\psi_j^\gamma l_{jk}^\gamma}\right)^{\frac{1}{1-\gamma}}\right].$$

Thus, defining the useful term Ω as

$$\Omega_n \equiv \frac{(1-s_n)Y_n}{\sum_k(\xi_{kn}^{-1}\psi_k l_{kn})^\gamma}, \tag{B.14}$$

we can rewrite (B.13) as

$$\frac{l_{in}}{l_i} = (\gamma\xi_{in}^{-\gamma}\psi_i^\gamma)^{\frac{1}{1-\gamma}} \times \Omega_n^{\frac{1}{1-\gamma}} \times \left[\sum_k\left[\left(\gamma\xi_{ik}^{-\gamma}\psi_i^\gamma\right)^{\frac{1}{1-\gamma}} \times \Omega_k^{\frac{1}{1-\gamma}}\right]\right]^{-1},$$

Which implies that optimal l_{in} is given by:

$$l_{in} = \frac{\xi_{in}^{\frac{-\gamma}{1-\gamma}} \Omega_n^{\frac{1}{1-\gamma}}}{\sum_k \xi_{ik}^{\frac{-\gamma}{1-\gamma}} \Omega_k^{\frac{1}{1-\gamma}}} \times l_i. \tag{B.15}$$

Plugging back (B.15) into (B.12), one gets equilibrium probability of contest winning, p_{in}^*:

$$p_{in}^* = \psi_i^\gamma \xi_{in}^{-\gamma} \times \left[\frac{\xi_{in}^{\frac{-\gamma}{1-\gamma}} \Omega_n^{\frac{1}{1-\gamma}}}{\sum_k \xi_{ik}^{\frac{-\gamma}{1-\gamma}} \Omega_k^{\frac{1}{1-\gamma}}} \times l_i \right]^\gamma \times \left[\sum_k \psi_k^\gamma l_{kn}^\gamma \xi_{kn}^{-\gamma} \right]^{-1}$$

2. The next step is to use equilibrium p_{in}^* in total looting revenues of i

$$R_i = \sum_n p_{in}^* \left(1 - s_n\right) Y_n. \tag{B.16}$$

Noting that ψ_i^γ, l_i^γ and $\sum_k \xi_{ik}^{\frac{-\gamma}{1-\gamma}} \Omega_k^{\frac{1}{1-\gamma}}$ do not depend on destination n:

$$\begin{aligned}
R_i &= (l_i \psi_i)^\gamma \left[\sum_k \xi_{ik}^{\frac{-\gamma}{1-\gamma}} \Omega_k^{\frac{1}{1-\gamma}} \right]^{-\gamma} \times \sum_n \frac{(1-s_n) Y_n \xi_{in}^{-\gamma}}{\sum_k \psi_k^\gamma l_{kn}^\gamma \xi_{kn}^{-\gamma}} \times \left(\xi_{in}^{\frac{-\gamma}{1-\gamma}} \times \Omega_n^{\frac{1}{1-\gamma}} \right)^\gamma \\
&= (l_i \psi_i)^\gamma \left[\sum_k \xi_{ik}^{\frac{-\gamma}{1-\gamma}} \Omega_k^{\frac{1}{1-\gamma}} \right]^{-\gamma} \times \sum_n \Omega_n \xi_{kn}^{-\gamma} \times \left(\xi_{in}^{\frac{-\gamma}{1-\gamma}} \times \Omega_n^{\frac{1}{1-\gamma}} \right)^\gamma \\
&= (l_i \psi_i)^\gamma \left[\sum_n \Omega_n^{\frac{1}{1-\gamma}} \xi_{in}^{\frac{-\gamma}{1-\gamma}} \right]^{1-\gamma}
\end{aligned}$$

Using the fact that looting revenues are redistributed among fighters, each paid w_i^F, we have $R_i = w_i^F l_i$ and we can solve for l_i:

$$l_i = \frac{\psi_i^{\frac{\gamma}{1-\gamma}}}{(w_i^F)^{\frac{1}{1-\gamma}}} \sum_n \Omega_n^{\frac{1}{1-\gamma}} \xi_{in}^{\frac{-\gamma}{1-\gamma}}. \tag{B.17}$$

This allows to simplify further the expression for the optimal l_{in} in (B.15) to obtain:

$$l_{in} = \frac{\psi_i^{\frac{\gamma}{1-\gamma}}}{(w_i^F)^{\frac{1}{1-\gamma}}} \xi_{in}^{\frac{-\gamma}{1-\gamma}} \Omega_n^{\frac{1}{1-\gamma}} \tag{B.18}$$

3. As can be seen from (B.14), the above equation still has numbers of fighters (l_{in}) on both sides through the Ω terms. The next step is therefore to solve for Ω_n as being functions of w rather than l. In order to do that, start with definition of Ω_n to obtain

$$\sum_k (\psi_k l_{kn} \xi_{kn}^{-1})^\gamma = \frac{(1-s_n) Y_n}{\Omega_n}.$$

Then replacing equilibrium bilateral allocation of fighters from (B.18) into $\sum_k (\psi_k l_{kn} \xi_{kn}^{-1})^\gamma$

yields

$$
\begin{aligned}
\sum_k (\psi_k l_{kn} \xi_{kn}^{-1})^{\gamma} &= \sum_k \left(\psi_k^{\frac{1}{1-\gamma}} (w_k^F)^{\frac{-1}{1-\gamma}} \xi_{kn}^{\frac{-1}{1-\gamma}} \Omega_n^{\frac{1}{1-\gamma}} \right)^{\gamma} \\
&= \Omega_n^{\frac{\gamma}{1-\gamma}} \sum_k \left(\frac{\psi_k}{w_k^F \xi_{kn}} \right)^{\frac{\gamma}{1-\gamma}} .
\end{aligned}
\tag{B.19}
$$

As a result, we obtain a solution for Ω_n which is independent of l_{in}:

$$
\Omega_n^{\frac{1}{1-\gamma}} = \frac{(1-s_n)Y_n}{\sum_k \left(\frac{\psi_k}{w_k^F \xi_{kn}} \right)^{\frac{\gamma}{1-\gamma}}}. \tag{B.20}
$$

Defining $\texttt{violence}_{\texttt{in}} \equiv \psi_i l_{in}$ and combining (B.18) with (B.20) gives the final equation for gravity of violence:

$$
\texttt{violence}_{\texttt{in}} \equiv \left(\frac{\psi_i}{w_i^F} \right)^{\frac{1}{1-\gamma}} \xi_{in}^{\frac{-\gamma}{1-\gamma}} \frac{(1-s_n)Y_n}{\sum_k \xi_{kn}^{\frac{-\gamma}{1-\gamma}} (\frac{\psi_k}{w_k^F})^{\frac{\gamma}{1-\gamma}}} \tag{B.21}
$$

□

Finally, we can note that the equilibrium success probability (using B.18 in B.12) is

$$
p_{in} = \frac{\left(\frac{\psi_i}{w_i^F \xi_{in}} \right)^{\frac{\gamma}{1-\gamma}}}{\sum_k \left(\frac{\psi_k}{w_k^F \xi_{kn}} \right)^{\frac{\gamma}{1-\gamma}}}. \tag{B.22}
$$

Therefore, after optimizing the allocation of fighters over space, we have that

$$
\texttt{violence}_{\texttt{in}} = \frac{\psi_i}{w_i^F} p_{in}(1-s_n)Y_n \Rightarrow w_i^F l_{in} = p_{in}(1-s_n)Y_n,
$$

which implies that the financial flow linked to bilateral violence simply multiplies lootable income with the probability of success.