



The Pearson Institute Discussion Paper

The Forging of a Rebel

Gauthier Marchais

Christian Mastaki Mugaruka

Raúl Sánchez De La Sierra

David Qihang Wu

2021-18



THE PEARSON INSTITUTE
FOR THE STUDY AND RESOLUTION OF GLOBAL CONFLICTS

THE FORGING OF A REBEL*

Gauthier Marchais Christian Mastaki Mugaruka
Raúl Sánchez de la Sierra David Qihang Wu[†]

April 18, 2021

Abstract

We use variation in exposure to victimization of 1,537 households of eastern Congo for each year of 1990–2013 to examine the formation of preferences to participate in armed groups. In this context, most armed groups are Congolese militia, whose objective is fighting foreign armed groups. We find that foreign armed group attacks on household members are associated with a larger propensity that individuals join a Congolese militia in subsequent years. The results are consistent with the formation of preferences arising from parochial altruism towards the family to fight foreign perpetrators. Specifically, we find that the effect is driven by the most gruesome of those attacks, by those that take place at a young age, and persists for several years. Consistent with parochial altruism, we find that the effect is largest when the victim is a household member or the village chief, smaller when the victim is another household in the village, and insignificant if the victim is in a nearby village. To examine the external validity of our result, we analyze heterogeneous effects by weakness of the state. We find that the response is concentrated in village-year observations in which state forces are absent. Finally, we show that, to undo this effect, the yearly per capita income outside armed groups would have to permanently increase 18.2-fold. These results suggest that intrinsic preferences are important for armed group participation relative to economic incentives, and emphasize their interaction with state weakness.

*The data collection was funded by Private Enterprise Development in Low-income countries (PEDL), the National Science Foundation, and the International Center for Taxation and Development (ICTD). We are thankful to Aimable Amani Lameke, Desire Basibuhe, Lewis Dunia Butinda, Emmanuel Til Kandate, Freddy Koleramungu, Eustache Kuliumbwa, Simeon Lukeno, and Christian Polepole Bazuzi for outstanding field research. Soeren Henn, Matthew Pecenco, and Carlos Schmidt-Padilla provided extraordinary research support. We thank James Robinson, Soeren Henn, Laura Montenegro, Jean Decety, and participants of Workshop on Conflict 2019 at the Paris School of Economics for comments. The title is a reference to Arturo Barea’s autobiography during the Spanish Civil War (Barea, 1984). Word count: 13,924 words.

[†]Affiliations: Institute of Development Studies, University of Sussex; Marakuja Kivu Research; Harris School of Public Policy, University of Chicago and National Bureau of Economic Research; Department of Economics, University of California, Berkeley

1 Introduction

War and violence are often socially costly.¹ Armed groups, of which popular militias, transnational violent organizations, and national armies are predominant forms, constitute a threat to existing states and societies across the world, both now and historically. These organizations are typically labor intensive, hence their success hinges on their ability to recruit.

In this paper, we study why individuals join armed groups. A predominant form of recruitment today is through voluntary entry. Yet, participation in these organizations entails life-threatening activities and violence against others. A puzzle to their success, then, is the following: why do individuals choose to join armed groups in the first place?

One possibility is that individuals join armed groups because they expect to receive a high income in the armed group. For this to be the dominant motive—given the high risk it entails and possible disutility of exercising violence—armed groups would probably need to offer high compensating differentials. Some scholars have shown that expectations of income matter in the context of gangs (Levitt and Venkatesh, 2000). Others have argued that economic incentives could play a role in civil conflict, although observing transitions into armed groups has been impossible (Collier and Hoeffler, 1998; Dube and Vargas, 2013). Despite some evidence that economic incentives have a significant effect on violent events, there is no evidence that they have much power to explain armed group recruitment.

Another possibility is that individuals join armed groups for nonmaterial motives. Armed groups, especially popular militia, typically have clear objectives, which may align with individuals' intrinsic motivations. José Manuel Mireles Valverde, one of the leaders of the Autodefensas movement in Michoacán, Mexico, explained that every single one of the members of his armed self-defence militia has lost a relative or close friend to the drug cartels. That experience, he explains, is the foundation of their commitment to the movement.² Many scholars in other social sciences have emphasized the importance of intrinsic motiva-

¹On the social costs of war, see Fearon (1995), Collier (1999), Hoeffler (2017).

²See Heineman (2015).

tion to explain armed group recruitment, and specifically militia recruitment. Economists have recognized that nonmaterial motives could be important for civil conflict (Collier and Hoeffler, 2004). However, there is no empirical evidence in economics about the importance of intrinsic motivation for armed group recruitment, except through self-reported narratives. This paper attempts to settle this debate and demonstrate how nonmaterial motives form.

In this paper, we provide evidence for the formation, and effects, of intrinsic preferences to join armed groups, in eastern Democratic Republic of the Congo (DRC). A number of characteristics of eastern DRC make the setting well suited for identifying the formation of preferences to join armed groups. First, today, around 120 nonstate armed groups operate in eastern DRC (Vogel et al., 2021). Second, some of these armed groups are perceived as foreigners. Indeed, several of these groups have origins in neighboring countries (henceforth, foreign armed groups). For instance, the Front de Liberation du Rwanda (FDLR) was established by former Rwandan military officers and militias involved in the Rwandan genocide (Stys, 2014; Johnson et al., 2016). The FDLR has engaged in considerable violence against civilian populations in eastern DRC. On the other side of the divide, numerous local militias have emerged to oppose foreign armed groups, often designated as Mai-Mai (henceforth, Congolese militia). Qualitative research has suggested that victimization of the individuals' family by foreign armed groups may be a leading explanation for why individuals join a Congolese militia (Marchais, 2016).

As a foundation for this study, we assembled a yearly panel dataset on the occupational choices and household histories of 1,537 households from 239 municipalities in North and South Kivu provinces, and the violence perpetrated by armed actors on those households, dating back to 1990. We measure exposure to attacks on the household using a cross-validation of various sources, which we constructed through one week of fieldwork by a team of two field researchers in each village. We measure participation into armed groups using household histories reported by each respondent. At the time of the data gathering, episodes of participation were either complete or in a district (Shabunda) controlled by a

local militia. Thus, everywhere in our context, with appropriate care that we describe in this paper, participation could be reconstructed. Our main analysis exploits variation in exposure to foreign armed group attacks across and within households over time.

We first compare the mean characteristics of individuals at the time when they join armed groups vs. the rest of individuals and years in the sample. We analyze 296 episodes in which an individual participates in an armed group. The mean duration is 2.34 years. We find that 78% of those correspond to participation in a Congolese militia. The other episodes are in either the Congolese army or in foreign armed groups. Yet, mean socioeconomic characteristics of participants and nonparticipants are almost identical, with one exception: the household of individuals who have participated in an armed group is 2.5 times more likely to have been attacked in the past by a foreign armed group. This could indicate that foreign armed group attacks on the family lead to joining, possibly a Congolese militia.

To isolate the causal effect of an attack by foreign armed groups, our analysis then compares individuals at years in which their households have not previously been attacked to years in which their households have previously been attacked, within year, individual, village, and cohort. Specifically, we regress an indicator for participation in a Congolese militia on an indicator for whether the household had previously been attacked by a foreign armed group, and include individual, year, village, and age fixed effects. We find that if the household has been attacked by a foreign armed group, the probability that the individual participates in a Congolese militia is 2.55 pp (2.36 times) larger in each subsequent year. This effect is so large that it drives the effect of attacks by any armed group on participation into any armed group. It is consistent with the formation of intrinsic motivation.

We then examine potential threats to inference: endogeneity of foreign armed group attack, nonclassical measurement error, selection bias through migration, persistence of participation, coincidental aggregate shocks, and whether foreign armed group attack captures coincidental conflict activity. Using the timing of historical events that affected a sub-set of households, observable characteristics of households attacked and not attacked, the individ-

uals' migration histories, the reports of attacks by other households in the same village, and analyzing bias in reporting, we find no evidence of confounding.

Our analysis then turns to the underlying mechanism. We examine three hypotheses that provide support to the formation of preferences to fight arising from parochial altruism.

First, if the finding reflects the formation of preferences, then it should arise from attacks that are more traumatic. We find that the effect is driven by the attacks that are the most gruesome and by attacks that took place when the respondent was at a young age.

Second, if the effect reflects the formation of preferences, it should be stable, and hence persist. We find that the effect persists beyond five years after the attack.

Third, we examine whether such preferences reflect empathic concern for one's group (henceforth, parochial altruism). Indeed, we find that, while attacks on the village chief, on other households, and on other villages cannot explain our result, the effect of a foreign armed group attack has an independent, but monotonically decreasing effect in the social distance of the victim to the respondent — i.e., it is largest when the victim is a household member or the village chief, it is smaller when the victim is another household in the village, and insignificant if the victim is another household of the Chiefdom outside the village.

Our analysis then turns to leading alternative causal channels that may explain our main effect. Guided by a household dynamic optimization model, we analyze whether our baseline estimate could reflect each of the following three alternative mechanisms. First, attacks on the household may have depleted the household's capital stock, altering the incentives for occupational choice and potentially causing participation into armed groups. Second, attacks on the household may increase the household members' concern for security, potentially leading them to join armed groups expecting to obtain private protection. Third, attacks on the household could have led the respondents to join armed groups because the respondent may be directly victimized, which could have reduced the productivity of their labor in the labor market. Using detailed household asset histories, the cross-validation of reported attacks in the village, and the details of who is victimized, we rule out these channels.

To assess the conditions for external validity of our result, we then examine whether our baseline estimate depends on the ability of the state to repress armed group participation. Specifically, we analyze the interaction between past victimization of the household and the current presence of state forces. Exploiting changes in the presence of state forces over time and across villages, and also using the timing of military redeployments as an instrument, we find that the entire behavioral response to foreign armed group attacks is driven by participation in a Congolese militia during years in which the state forces are absent—often those take place various years after the attack. This suggests that, while victimization creates persistent preferences to fight back foreign armed groups, those are channeled into participation in a Congolese militia if, and only if, the state is weak.

Our analysis then turns to benchmarking the magnitude of our estimate to the effect of the withdrawal of state forces, and to the effect of income fluctuations. First, we find that the magnitude of our effect is comparable to the effect of the withdrawal of state forces. This suggests that state efforts to repress armed groups could be replaced by state efforts to protect civilians. Second, we find that to undo the effect of an attack, the yearly revenue outside armed groups would need to increase, each year, 18.2 fold. To establish this result, we exploit a large shock to the world price for gold. Gold is mined by individuals in some locations of our sample, but not others. Gold is very easy to conceal. Thus, its price increases income in mining more than it does that that can be obtained by joining armed group.

Our study complements a burgeoning literature in economics that examines the motives for violence (Blattman and Miguel, 2010). Social scientists have long discussed the role of nonmaterial motives (Kalyvas, 2006; Wood, 2006; Arjona, 2008; Balcells, 2012).³ While economists recognize the possibility of nonmaterial motivations (Collier and Hoeffler, 2004),

³The literature in political science, sociology and anthropology identifies several factors that explain participation in violence. These include grievances resulting from relative deprivation (Gurr, 1970), selective monetary incentives (Popkin, 1979; Weinstein, 2007), social networks that mediate recruitment (Gould, 1993; Parkinson, 2013; Staniland, 2014; Scacco, 2016), coercion and the cost of nonparticipation (Kalyvas and Kocher, 2007), and ideology (Gutiérrez Sanín and Wood, 2014; Atran, 2016; Schubiger and Zelina, 2017). A consensus is that various factors are relevant (Humphreys and Weinstein, 2008; Marchais, 2016; Viterna, 2006, 2013; Parkinson, 2013). However, recent contributions in political science emphasize the role of moral beliefs and emotions (Wood, 2003) and of victimization and revenge (Balcells, 2017).

to date, the study of civil war in economics remains focused on its economic logic (Fearon, 1995; Jackson and Morelli, 2011; Dube and Vargas, 2013; Sánchez de la Sierra, 2020).

The findings also contribute to the study of the effects of war on pro-social behavior (Bauer et al., 2016). Due to the difficulty of observing data on war, war exposure remains a “black box,” and pro-social behavior is typically measured in the lab or in the form of peaceful political participation. Complementing this literature, we micro-found the channels through which war leads to altruistic behavior that is privately costly in a real context, and characterize how it forms, and how the institutional environment mediates its expression. When the state is “weak,” its expression is notably through joining Congolese militia.

Our study also complements the study of the effect of coercion and violence on development. Seminal research has examined the effects of coercion motivated by extraction of resources on wealth, education, health, social cohesion, and institutions (Lowe and Montero, 2021). Violent practices by the colonial state are of similar degrading nature to those perpetrated by the foreign armed groups in our context.⁴ Complementing the findings in this literature, we show that individuals have agency despite these experiences, and use it to combat the causes of oppression. Our findings suggest that individuals use the channels that are available in their context to express their newly seeded preferences for rebellion.

Finally, this paper also complements the literature on intrinsic motivation (Bénabou and Tirole, 2003). Seminal studies have analyzed the margin of selection of workers with intrinsic motivation into organizations with a public good-oriented mission (Dal Bó et al., 2013; Ashraf et al., 2014a,b; Dal Bó et al., 2017; Deserranno, 2019; Karing, 2021). Due to the difficulty to measure its origin, intrinsic motivation is often taken as exogenous, or experimentally re-activated (Khan, 2021). We provide evidence for how it can form at an early age, and show that it plays a role at a high-stakes developmental outcome.

We now turn to a discussion of the Congolese militia in our analysis.

⁴For instance, Lowe and Montero (2021) quote: “My father was murdered: they tied him to a tree and shot and killed him, and when the sentries untied him they gave him to their boys, who ate him. My mother and I were taken prisoner. The sentries cut off my mother’s hands while she was still alive.”

2 Context

Three characteristics make the provinces of North and South Kivu in the last 30 years well suited to study the formation of intrinsic motivation to fight.

First, since the First (1996–97) and Second (1998–2003) Congolese Wars, rebel factions from Rwanda and Burundi have consolidated bases in rural areas of North and South Kivu (henceforth, foreign armed groups). The narratives of the population suggest that the violence of foreign rebel factions have motivated many Congolese to take up arms.

Second, this conflict taps onto preexisting divisions and distrust between so-called “autochthonous” communities and “foreigners,” in particular Rwandophone armed actors (Vlassenroot, 2002; Jackson, 2006; Vogel and Stearns, 2018; Sánchez de la Sierra, 2021). These divisions date back to the 19th century and have been reinforced by the violent conflicts of the 1960’s and 1990’s (Hoffmann and Vlassenroot, 2014; Hoffmann et al., 2017).

Third, the 1990’s have seen a proliferation of armed factions in the region, from village self-defence militia to larger armed groups. Their proliferation is the result of international military confrontations, conflicts over land, resources and customary authority, integration and defection from the Congolese national army, and protection of civilians.⁵ Today, there are around 120 of them in the eastern provinces of the DRC (Vogel et al., 2021). Among these, many are so-called Mai-Mai militia. The Mai-Mai are a constellation of armed groups, originally formed during the Congolese wars as part of a large-scale armed resistance movement, which splintered into many smaller groups. Mai-Mai factions typically have a strong social basis in the rural areas of South Kivu (Hoffmann and Verweijen, 2018). In rural areas, joining the Mai-Mai has become part of life for many Congolese, often young men. Qualitative research suggests that, among various reasons for joining, motives include the desire to fight foreign armed groups who have abused relatives or the community.⁶

⁵For an in depth analysis of the causes of armed mobilization in eastern DRC, see The Rift Valley Institute’s Usalama Project I: Understanding Congolese Armed Groups and II: Governance in Conflict, in particular Stearns (2012b); Stearns et al. (2013); Vlassenroot (2013); Verweijen (2016).

⁶Other motives include revenge, social exclusion, income, protection, ideology, and, in some cases, coercion (Vlassenroot and Van Acker, 2001; Jourdan, 2004, 2011; Brabant, 2016; Marchais, 2016).

The case of the Raia Mutomboki (“outraged citizens”) illustrates the role of parochial altruism to create intrinsic motivation to join militia. At the end of the Second Congolese War (1998-2003), most of the armed factions who had been operating in the rural areas of South Kivu were incorporated into the new Congolese National Army, the Forces Armées de la République Démocratique du Congo (FARDC). Their withdrawal created a security vacuum in the territory of Shabunda, South Kivu: indeed, factions of the Front de Libération du Rwanda (FDLR), a Rwandan armed group with origins in the Rwandan Civil War, increased human rights violations against civilians in Shabunda. In response, a local Kimbanguist Minister, Jean Musumbu, mobilized the population to fight the FDLR aided by spiritual powers and magical amulets known as *dawa*. This was the emergence of the Raia Mutomboki, which garnered wide popular support and drove the FDLR out of southern Shabunda (Stearns, 2013; Vogel, 2014). It reemerged in 2011 under similar circumstances, and acquired a larger scale (Stearns, 2013; Hoffmann and Vlassenroot, 2014; Vogel, 2014).

The Raia Mutomboki tapped onto grievances created by the FDLR’s violence against the population. In the initial stages of the 2004–05 and 2011 mobilizations, the movement enjoyed considerable popular support, displaying the characteristics of a social movement (Stearns and Botiveau, 2013). The decentralized character of the movement allowed franchise-like local chapters to emerge in villages and towns of Shabunda (Stearns, 2013). Participants were typically free to enter and leave. The movement’s message also articulated grievances of the population against neglect by the Congolese state and their suffering at the hands of the FDLR with a call to action to rid the area of the FDLR.⁷

Qualitative interviews carried out for this project provide evidence of widespread abuses by the FDLR, and of these abuses being a driver of Congolese militia mobilization. Consider the following account by three notables in the village of Nyambembe:

[The FDLR] acted as mercenaries, pillaging and carrying out massacres on the population . . . They became more and more harmful until the population

⁷Like many other armed groups in the DRC (Hoffmann, 2015), the movement’s use of the mystico-religious beliefs of the Rega further strengthened its popular appeal (Stearns, 2013).

organized themselves, given that the national army was incapable of stopping the attacks of the FDLR . . . The Raia Mutomboki emerged in this village in 2011, and was essentially composed of *autochtones*, with the objective of chasing away the [FDLR] . . . At first, they reacted when they learned that the FDLR were in the village, but with the increase in FDLR violence they started attacking them directly, and thanks to their *fetiches*, they were able to get some weapons from the FDLR and chased the FDLR out of Shabunda. (Interview carried out in September 2013 in Nyambembe).

Consider the following account from a trader of Misau village, in Walikale district:

[The FDLR] forced me to leave the bed and asked me to hand over everything I had as money, they emptied my shop. They even forced my wife to sleep with them in my presence, and after pillaging our shop, they shot bullets in the air, pillaged the whole village, and burnt some households before leaving. A day after these events, the national army came to secure the village upon the request of the population. They organized patrols regularly to eradicate the FDLR, and the youth of the village joined the army to help them fight the FDLR, including myself (phone interview, March 2021 in Walikale).

Consider the following account from a trader of Pinga village, in Walikale district:

From afar, I saw the body of my father and two of my brothers already with their throats slit, and I saw the [FDLR], I ran to the village to inform the Mai-Mai NDC. They gave me the grigri to protect me against bullets, and upon returning to the scene, the FDLR had left, I fainted when I saw the bodies of my family. We had to bury them in the village. After this experience, the NDC began to organize patrols to track the FDLR and, since I knew the area, I helped their operations and later joined the NDC (phone interview, March 2021 in Walikale).

Section B provides additional background on the FDLR and Congolese militia.

3 Data

Our sample comprises 1,537 households living in a random sample of villages of South Kivu and North Kivu, eastern provinces of the DRC.⁸ For each household, we observe the history of exposure to violence and asset ownership and investment for the household, and the history of participation in armed militia for one of its members.

Participation in militia is commonplace in eastern DRC, hence discussing episodes of participation is conceivable. Many armed groups such as the Mai-Mai or Raia Mutomboki have enjoyed support from the population. As a result, it is possible to discuss past and current episodes of participation and with appropriate measures, do so without fear of retribution or social desirability bias. Thus, we are confident that motives for concealing participation were sufficiently contained that they do not affect the validity of our exercise.

Prior to gathering information, we obtained authorizations from provincial, territory, and village state authorities, as well as customary authorities to whom the project was presented in detail. Ethical guidelines were followed to ensure that respondents did not feel obliged to participate in the study and that their participation didn't expose them or the researchers to any risk. Field researchers spent at least a week in each village or entity, during which they built relationships with respondents and local communities, and gathered information on the histories of the entities.

The section that follows describes the data collection.

3.1 Data collection

The data were collected between June 2012 and September 2013 in South Kivu, and between June 2015 and June 2016 in North Kivu. Because no census has been conducted in recent years, in a first data-gathering round, members of the research team spent weeks in the districts' (so-called Chiefdoms) capitals and in the lower-level districts (so-called groupements)

⁸We exclude villages without valuable resources from the sampling frame, as those see typically lower levels of armed group activity. See Sánchez de la Sierra (2020) for more details on data collection.

to draw lists of all villages in each district by consulting state and customary authorities. In those lists, we identified the villages that had a natural resource—the rest typically had less armed group activity. From that list, we randomly sampled 133 villages in South Kivu and 106 villages in North Kivu.

Once we had consolidated the sample of villages, teams of two researchers visited each village. In each village, the assigned team worked for one week, reconstructing the history of the village, and of a random sample of households. Since there was no census, they first drew a village list with the village chief, and implemented random selection using pre-selected random numbers generated through a computer. In South Kivu, we randomly sampled eight households in each village. In North Kivu, we randomly sampled six households per village.⁹

At the same time, the researchers conducted a day-long household survey in each sampled household. In each household, they randomly sampled one available male adult to work with the researcher during one full day, with appropriate compensation and food. The survey comprised breaks for lunch, informal conversations, ethnographic in-depth data gathering, and reconstructed the household’s and the respondent’s history, dating back to at least 1995.

Our sample comprised 1,537 households (1,041 in South Kivu and 496 in North Kivu). The survey gathered yearly information about the respondents’ history of participation in armed groups, detailed information of the nature and the dates of all violent attacks experienced by the household and any of its members, as well as a history of occupational choice, migration, and acquisition and liquidation of assets (cattle), but also about the village history of attacks and armed groups’ presence. This allows us to cross-validate the respondent reports with the information provided by other households and by the village history.

Reconstructing a household’s history based on recall is subject to classic measurement error of magnitudes, of event dates, and to nonclassical error due to sensitive information. To address these challenges, we used established methods in recall studies from eyewitnesses

⁹The field researchers lived with the community during one week. They built ties with the community, collected qualitative information about the history of the conflict in the village, worked with the history experts of the village to reconstruct, and verify the village history.

and, based on three months of piloting, tailored them to the cultural context.¹⁰ Specifically, we implemented the following measures.

First, in each province and for each year, we identified time cues that respondents would remember from their experience of the regional history. When asking a question about a historical event, such as an attack or the acquisition of cattle, to identify the year in which the event took place, the field researchers first examined whether the event was before or after the time cues in that area. Respondents sometimes did not know the exact year, but they recalled with certainty whether it was after or before a given time cue. Since we know the year of the time cues, this allowed the field researchers to pin down the exact years.

Second, built-in in the surveys was a strategy for auto-generation of person-specific relevant time-cues, anchored to common knowledge historical time cues. For instance, at the start of each survey, the field researcher asked about the easiest information to recall: when they were born, when they got married, when they migrated (if applicable), using the historical common knowledge time cues. These life events provided respondent-specific time cues that field researchers then were trained to use for the remainder of the survey. This made it straightforward to determine the years at which the following events discussed in the survey took place, even when respondents were not sure a priori about the year.¹¹

Third, we also administered working memory measurements. This allows us to weight observations by the ability of the respondent to memorize numbers, for robustness.

3.2 Description of survey measures

Based on our own qualitative interviews (Marchais, 2016; Sánchez de la Sierra, 2020) and extensive discussions with specialists of the survey regions, we classify all the armed groups into: (1) Congolese militia, (2) Foreign armed groups, (3) Congolese army.¹²

¹⁰Sánchez de la Sierra (2020) presents additional information about the method for this data collection.

¹¹For instance, it was always easy for the respondent to answer whether they had acquired a cow before or after their marriage, or before or after the Second Congo War started.

¹²We built this classification on the basis of existing literature on the eastern Congolese conflict, qualitative fieldwork, and our quantitative data. For each armed group reported in the survey, we collected information

Victimization. The sources are a household attack history module and a household roster.

In the household attack history module, each respondent is asked to report up to nine attacks by armed actors that happened in the village where they live in. Each respondent on average reports 2.08 attack events; the 99%th percentile is seven events. Thus, reporting limit did not lead to loss of data. For each event, we obtain perpetrators' identity, the perceived intention behind the attack, whether respondent was physically victimized, whether the household was pillaged, the number of fatalities in the village, the number of persons who suffered sexual violence in the village, and whether the village chief was victimized.¹³

The household roster contains information on all the members of the household at the time of the survey. For each household member, each respondent reported (1) up to three events in which armed actors victimized the household member directly, and (2) of these, up to three events in which the armed actor perpetrated sexual violence on the household member. For each of those events, we identify the year in which they took place.¹⁴ Based on this information, we construct an indicator for whether respondent reported an attack on the household members for each year. We link the attack information in both modules.¹⁵

In the analysis that follows, to isolate parochially altruistic motivations, we focus on whether any household member other than the respondent was victimized, henceforth household victimization. This includes episodes in which the respondent was also victimized.

Participation in armed groups. We obtain information on whether the respondent participated in militias or armed groups.¹⁶ The information comes from the security module in

on the name of the group, whether the group had been formed in the survey entity, the nationality of the leaders of the group and whether the group was perceived as local to the entity. Given the franchise character of armed movements such as the Mai-Mai or Raia Mutomboki, composed of numerous subgroups, the questions were geared toward the group that was present in the entity and not the movement as a whole. Furthermore, we triangulate the classifications of the household survey with those of the village survey.

¹³We only ask whether village chief is victimized in the South Kivu survey.

¹⁴Since the household information module also elicits whether respondents themselves are attacked, this information overlaps with respondent attack module where respondents are asked whether they suffer from physical assault during attacks on village. We clean the latter information using the former.

¹⁵We use the attack information from both modules for each household member attack. See Appendix A.

¹⁶For measurement purposes, we define participation as the active involvement in the security-related activities of an armed group. There are other forms of involvement: civilians can work as informants, covert supporters, tax collectors, enforcers, business partners, and other roles. Our definition of participation does not include these other forms of involvement. In Petersen (2001)'s classification, our definition corresponds to

the household survey. To obtain this, we first ask each respondent to list the armed groups that have been in the village where they lived at the time of the interview. For each armed group, each respondent is asked whether he had participated in it and if yes, the start date and end date. In addition, for respondents not born in the village of interview, in a separate module, we asked the respondent to describe each episode of participation in an armed group in years preceding the arrival to this village. Based on this information, we construct the respondent’s history of participation in a Congolese militia (henceforth, “participation”).

The information on victimization was gathered prior to that about participation in armed groups. This design feature was intentional. It helped shield the data gathering against motivated reasoning biasing the reporting of attacks. For instance, if a respondent reported to have participated he may then be tempted to overreport or provide more effort to remember the attack events, as a way to rationalize his decision to participate. By asking about the attacks well before asking about participation, this source of bias is made impossible.

Household economic history. We also observe the household’s yearly investment and occupational history.¹⁷ For each year since 1995, we observe how many cows, goats, and pigs the household bought or sold; how many fields the household bought or sold; whether respondent held a marriage; and whether the respondent worked in any of the following categories: agriculture, school student, mining, government-related jobs, or unemployment.

Village history. We obtain the history of armed group presence in the village to construct an indicator for whether there is a state force present in the village in a given year (henceforth, “weak state” indicator). Given the context history, we consider as state forces the Congolese army (for up to 1996 and after 2004) and the Rassemblement Congolais pour la Démocratie (RCD) (for 1998–2004), who had taken over the apparatus of the state during the Second Congo War (Stearns, 2012a). We also use this information to cross-validate household attack information with village attack information from chief and village specialists (see Section 6.2).

the second and third levels of involvement in insurgent groups, both the active participation in local security organizations (the second level) and participation in mobile armed factions (third level). Our measure of participation was extensively discussed with the research team in qualitative fieldwork.

¹⁷Information of farm animals is only asked in the South Kivu survey.

4 Characteristics of participation and attacks

There are 25,091 household-year observations in which the respondent is above age 15, spanning the years 1995 to 2013. In 3.1% of those, the respondents participates in an armed group, and in 9.3%, his household members have previously been victimized, reaching 16.14% of the 1,537 respondents by the end of the sample period. At that point, 14.05% of respondents have participated in an armed group. Figure I presents the village sample in 2012.

We first present some basic characteristics of the attacks in the sample. Figure II shows characteristics of the raw attacks dataset. In total, there are 3,473 individual-year observations between year 1995 and 2013 in which the respondent reports an attack that takes place in the village they live in. Of those, 475 are reports of an attack incident in which household members, other than the respondent, were directly victimized (we call this household victimization). Panel A shows that 333 of those 475 household victimization attacks were perpetrated by foreign armed groups (70% of the 475 attacks). Those were mostly perpetrated by the FDLR and other foreign armed groups. Ninety-five of the 475 household victimization attacks were perpetrated by Congolese militia. Eighteen of the 475 events were perpetrated by the Congolese army.¹⁸ Panel B shows that the respondents' household members are victimized in 1.1% of individual-year observations (475 events), other households (among the households of the study sample) in the same village are also victimized in 4.3% of observations (1,836 events), and other villages (among the villages in our sample of survey villages) of the same Chiefdom are also victimized in 74% of observations (21,732 events).¹⁹ Panel C shows the distribution of attacks and crimes perpetrated during those attacks. We find that perpetrators directly victimize the spouse of the respondent, or his children, in 332 and 106 attacks out of the 475 household victimization attacks, respectively. We find that out of the 475 household victimization attacks, 185 included sexual violence against a

¹⁸Figure D.1 in the online appendix shows the distribution of attack perpetrators in detail. The figure also includes attacks perpetrated by the Congolese national army. Since the sample of Congolese army attacks is only 18, in the analysis that follows, we analyze attacks by the Congolese army only in Appendix Table C.1.

¹⁹There are 21 Chiefdoms in the sample.

household member. Of those, 126 included sexual violence against the respondent’s spouse. Finally, of the 475 household victimization attacks, 420 included theft of household property. In sum, the attacks that we consider in the analysis that follows are, on average, gruesome, often targeted the spouse of the respondent, and almost always included theft.²⁰

We then present basic characteristics of observations in which an individual participates in a Congolese militia, which we call participation, and those in which they do not. Baseline difference between individual-year observations in which respondents participate in an armed group and those in which they do not are shown in Table I.

Column “Not Part.” shows the characteristics of all individual-year observations in which the respondent does not participate in any armed group. Column “Start Part.” includes only the characteristics in the first year of participation.²¹ The table reports, for each individual-year observation, characteristics of participation episodes (Panel A), respondent socio-demographic outcomes (Panel B), household economic outcomes (Panel C), and exposure to attacks (Panel D). Symbols *, **, *** indicate whether the difference between the two columns is statistically significant at the 1, 5, and 10% levels, respectively.

In Panel A, the unit of observation is the individual participation episode. The panel presents characteristics of the 296 episodes.²² Of the episodes, 78.4% were in a Congolese militia, 11.2% in foreign armed groups, and 2.03% (n=6) in the Congolese national army. The average duration of an episode is 2.34 years. The 296 episodes represent 251 individuals.²³

Panel B presents mean socio-demographic characteristics of the individuals for participation and nonparticipation episodes. The mean characteristics across these two groups are fairly balanced: individuals in the first year of participation episodes are not more likely to be married, to previously work in the mining sector, agriculture, as a civil servant, or to be unemployed. In the year in which they initiate a participation episode, individuals who participate tend to be one year younger than the rest of individuals in the sample.

²⁰Figure D.2 shows the distribution of attack perpetrators and motives in detail.

²¹We exclude 488 observations that are after the first year of the participation episode.

²²Figure D.3 presents the average duration of participation by year of episode start.

²³Due to its small sample, recruitment by the Congolese army is analyzed in Appendix Table C.1.

Panel C presents the household economic characteristics. Wealth at birth of individuals observed in participation episodes is comparable to that in nonparticipation individual-years. Their asset stock in the year before, and the investment in assets the subsequent year, is also comparable to that observed in individual-year observations not participating.

Finally, Panel D compares the exposure to past attacks. While participants tend to be indistinguishable from nonparticipants in other characteristics, individuals at the first year of a participation episode are 2.5 times more likely to have experienced a foreign armed group attack against their household members than the rest. This could indicate that foreign armed group attacks on the households cause individuals to participate in a Congolese militia.²⁴

5 Baseline estimate

We now turn to formally testing whether attacks by foreign armed groups on the household are associated with a higher propensity that individuals later join Congolese militia. Specifically, we estimate the following equation using ordinary least squares (OLS):

$$Part_{ijt} = \alpha + \gamma I_{it}^{Victim} + \alpha_i + \alpha_j + \alpha_t + \alpha_a + \mathbf{X}'_{it}\Gamma + \epsilon_{ijt} \quad (1)$$

where i, j, t index, individuals, villages, years, respectively. I_{it}^{Victim} is an indicator variable that equals 1 if other members in the household of individual i have been attacked at any period $t' < t$, and zero otherwise. $\alpha_i, \alpha_j, \alpha_t, \alpha_a$ are fixed effects for individual, village, year, and cohort (age), respectively.²⁵ The vector of individual-year level covariates, \mathbf{X}_{it} , is included where indicated. To account for serial correlation and village-level shocks, standard errors are two-way clustered at the individual and at the village*year (respectively, 1,537 and 9,061 clusters). Where indicated, we also estimate standard errors accounting for intra-Chiefdom correlation over space and time, i.e. clustering at the Chiefdom level (in that case,

²⁴Figure D.4 presents the episodes of participation (in blue) and the foreign armed groups' attacks on the household (red crosses) against years (x axis). Episodes of participation occur through the entire period.

²⁵All villages contain individuals who are observed in another village in some year.

since there are 21 Chiefdoms, we estimate the standard errors with wild bootstrap).²⁶

Estimates of equation 1 are reported in Table II. Columns (1)–(3) present separate regressions in which the dependent variable, $Part_{ijt}$, measures participation: in a Congolese militia, in a foreign armed group, in any armed group, respectively. The rows report the regressions in which I_{it}^{Victim} measures that the household members have been attacked by: a foreign armed group (By foreign group), a Congolese militia (By Congolese militia), any armed group (By any armed group). The table reports estimates (and standard errors) of γ .

The coefficient estimate in column (1), labeled “Congolese militia,” and row *By foreign armed group*, shows that individuals in households who have previously been attacked by a foreign armed group are 2.55 pp. more likely to participate in a Congolese militia. Since the control mean of participation in a Congolese militia is 1.88 pp, a foreign armed group attack on the household is associated with an average per year 136% larger propensity to participate in a Congolese militia. We refer to this relationship as the baseline estimate.²⁷ Its magnitude is so large that it drives entirely the effect of attacks by any armed group on participation in any armed group. Indeed, Column (3), labeled “Any armed group,” at row labeled *Any armed group* shows that an attack on the household by any armed group is associated with 1.73 pp. higher propensity to participate in any armed group, also significant.

In sum, this analysis shows that if the household has previously been victimized by a foreign armed group, the probability to join a Congolese militia is significantly larger. This suggests that foreign armed group attacks on the household cause participation into Congolese militia. However, various threats to identification challenge this causal interpretation. We analyze the merit of those threats in the next Section.²⁸

²⁶The number of villages included in the regression is 634, more than the 239 sample villages. It is because respondents might have lived outside the sample villages at a given year. We collect the migration history of each respondent which allows us to control for these out-of-sample villages in the regression.

²⁷Figure D.5 shows the main result graphically. This indicator is unchanged by whether the household members experience any other attack after the first one. However, 67% of the attacks in our sample are the first attack the household experiences in the household. The number of additional attacks is very small (only 2% of observations), even if an additional attack continues to have an independent, but insignificant, effect. Figure D.6 shows the baseline estimate broken down by first, second, and third attack. It is only significant in the first, consistent with the interpretation that it arises from trauma.

²⁸In the sections that follow, the analysis includes the data from both North and South Kivu provinces.

6 Examining potential confounders

In this section, we assess the merits of leading threats to validity of our baseline estimate.

6.1 Endogeneity of foreign armed group attack

A crucial identification assumption for interpreting the baseline estimate as quantifying a causal relationship is that the switch in victimized status within household across years is orthogonal to variables that are also associated with the propensity to join Congolese militia.

Motivated by this possibility, we checked for differences in observable characteristics between the household-year observations in which the household is attacked and those in which it is not.²⁹ We find that attacks do not disproportionately affect households with past participation, nor with different socio-demographic characteristics. However, attacks disproportionately affect more the wealthiest households. We found that this imbalance arises exclusively from attacks driven by pillaging. This suggests that victimized households may be self-selected, mostly due to by attacks motivated by pillaging. Reassuringly, we also found that richer households are not more likely to participate on average. We now analyze whether the baseline estimate is robust to selection of households into victimization.³⁰

Table III, Panel A, presents the analysis of endogeneity of past attacks on the household. Panel A.1 shows estimates of equation 1 using controls for selection. Column (1) estimates equation 1 as benchmark. Column (2) uses only foreign armed group attacks in which no pillage took place in the household as explanatory variable. Since, for attacks *not* motivated by pillage, households are balanced on observable characteristics as we show in Table C.2, estimation using this variable is less likely to be biased due to selection of households. The baseline estimate is strengthened by focusing on attacks not motivated by pillage.

Column (3) controls for all imbalanced variables presented in Table C.2. Time-invariant

All results arise from South Kivu respondents.

²⁹Table C.2 in the online appendix presents this balance.

³⁰Columns (1) and (2) of Table C.2 indicate that attack affected wealthier households disproportionately. Columns (3) and (4) decompose attacks into attacks in which attacks with the intention to pillage took place and the rest. We find that this imbalance arises exclusively from attacks driven by pillaging.

variables (wealth at birth) are interacted with indicators for years. We find that the coefficient is significantly larger, and statistical significance is strengthened. We also find that lagged asset stock is uncorrelated with participation, confirming that richer households are no more likely to participate in a Congolese militia. Accounting for these controls, a foreign armed group attack in the household is associated to a 178% larger future propensity to participate in a Congolese militia. The analysis in columns (2) and (3) suggests that selection of households cannot explain the baseline estimate, and thus is unlikely to be a confound.

Panel A.2 implements three different instrumental variable approaches that exploit plausibly exogenous sources of variation for attacks. Column (1) is the benchmark. In column (2), we instrument whether the household has been attacked in year t with whether the household lived in Shabunda during 1998–2002 or 2010–11, for two reasons. First, 1998–2002 is the height of the Second Congo war, the most intense period of the conflict, and the district of Shabunda, which was particularly affected by the conflict. Second, in 2011, the Congolese army retreated from rural areas of Shabunda for an army restructuring process known as Regimentation (Stearns, 2013). This led to a disproportionate rise in violent incursions by the FDLR. These are plausibly exogenous sources of exposure to attacks across villages and years. The estimated IV coefficient is large and significant, but the first stage is weak, which explains why the coefficient is very large (F-stat 3.96). Column (3) instruments household victimization by whether the village was controlled by a foreign armed group in 1998–2002 or 2010–11. The first stage is strong (F-stat 19.79) and the coefficient is positive and significant. This provides reassurance that selection of villages into household victimization does not explain the baseline estimate. Column (4) combines this village-year level variation with within village variation to instrument for household victimization within village. Specifically, column (4) uses the same instrument as column (3), interacted with an indicator for whether the household originates from a rich family. The coefficient is large and significant and the first stage is marginally strong (F-stat 8.34).

In sum, we find no evidence that selection of attacks can explain the baseline estimate.

6.2 Nonclassical measurement error

Another empirical concern is measurement error of participation and of attacks. For instance, respondents may have imperfect recall about participation, or about household victimization. The latter would imply that our baseline estimate is an under-estimate of the effect of household victimization. More severely, respondents could simply have systematically mis-reported victimization or participation. This could confound the baseline estimate in any direction. For instance, while the design of our survey makes this unlikely (see Section 3), past participants may have a tendency to remember better past attack events. We now formally assess the role of measurement error in explaining our baseline estimate.

In Table III, Panel B, we examine the role of measurement error. Column (1) presents the estimates for equation 1 as benchmark.

Column (2) estimates equation 1 using cross-validated attacks.³¹ For each attack reported by the respondent, we cross-validate the attack by looking at whether more than half of the other respondents who live in the same village in the same year also reported an attack on the village, or whether the chief reported an attack on the village in the same year: of the 475 attack episodes, 34.5% can be verified in this way.³² The coefficient is larger and significant.

Columns (3) and (4) estimate equation 1 weighting observations by their accuracy.³³ In column (3), for each respondent, we calculate the report precision by looking at the proportion of all the “true” attacks, defined as attack episodes reported by more than half of other contemporary respondents in the same village or by the village chief.³⁴ The coefficient magnitude is larger, and statistical significance above conventional levels remains. In column (4), we weight observations by the respondents’ score in the working memory tasks, assigning

³¹Figure D.9 shows the distribution of cross-validated attacks.

³²Many non-cross validated attacks happen in the same village but in a different year, indicating that they are measurement error of the year rather than mis-reporting. Among observations where respondents live in the current village, 126 out of 260 (48%) of the attacks are cross-validated. Only 20 out of 192 attacks where respondents live elsewhere can be cross-validated, since they report about attacks in other locations. For attacks in the in-survey villages, 142 out of 332 attacks can be verified in the chief survey.

³³Figure D.10 shows the distribution of precision measures.

³⁴For respondents who never experienced any “true” attack, we assign the mean of report precision.

a larger weight to observations from respondents with high scores in the working memory tasks. This discounts observations from respondents with poor working memory scores, who are arguably those providing the data with the highest likelihood of containing measurement error. The coefficient remains significant (and doubles) regardless which weights we use.³⁵

Column (5) implements an IV strategy to address classical measurement error by the respondents reports. Specifically, we use our village-level information for whether a foreign armed group attacked the village as an instrument for the respondents' report that their households were attacked in the village. With the IV strategy, the coefficient is larger (36.84%) and remains statistically significant at conventional significance levels.

The analysis in columns (2)–(5) thus suggests that the true effect is likely to be larger than the baseline estimate, this analysis thus confirms the validity of our conclusions.³⁶

6.3 Migration

An important assumption of our analysis is that migration is not a confound. We take advantage of the fact that we observe individuals over multiple years in the period. This allows us to include individual fixed effects, ruling out selection of individuals as a threat to internal validity: we estimate the treatment effect of an attack for the households present today in the sample of villages we interviewed. But migration can affect external validity of our results if the selection of households present today is not representative of those who were present in the past. In that case, we estimate the effect for a population subset.

Motivated by this possibility, we examine migration. Appendix G shows that coefficient γ of equation 1, when estimated in the sample of respondents who never leave the interview villages, is an unbiased estimator of the mean effect of household victimization in that sample.

³⁵Memory tests were only conducted in South Kivu. We constructed an index for accuracy using two tests. In a first test, the respondent is asked to repeat a predetermined sequence of randomly drawn digits multiple times, at increasing length of the sequence. For instance, if the sequence is 5293746, the surveyor first asks to repeat 5, then 52, then 529, etc., until the respondent fails to remember. The surveyor notes the first length at which the respondent failed to remember the sequence fully. In a second test, we asked them to count how many years had passed since Mobutu was no longer president.

³⁶Table C.3 implements an IV strategy using ACLED data.

Table III, Panel C, shows the analysis. Column (1) shows the baseline estimate as a benchmark. Column (2) estimates equation 1, but conditions the sample on villagers who never leave the sample of interview villages. The coefficient is 4.02, almost twice as large than the baseline estimate, and is significant. Column (3) estimates equation 1 in the full sample, but controls for whether the respondent lived outside the interview villages for at least one year in the study period. Similarly, the coefficient is doubled and remains significant.

In sum, we obtain a similar coefficient estimate with the population of households who never left the 239 interview villages. This corresponds to 1,086 out of 1,537 respondents in the sample, suggesting that, for such well-defined population, in addition to being unbiased, our estimate is also a close estimate for the households living today in the interview villages. This provides reassurance that migration does not affect the external validity of our analysis.

6.4 Other possible confounds

We now examine additional leading confounds due to mis-specification. Table III, Panel D, presents this analysis. Column (1), which presents the baseline estimate, is the benchmark.

We first account for the possibility that, since the average participation episode lasts 2.34 years, persistence of participation may confound the analysis. For instance, once an individual joins, the cost of leaving could be high. If this were the case, the baseline estimate may overestimate the effect of household victimization on the probability to join. Column (2) presents the estimates of equation 1, but in which the dependent variable takes value 1 only if the observation is the first year of a participation episode, zero otherwise. If the respondent participated twice, the first year for each episode takes value 1, the remaining years take value 0. The coefficient remains large and statistically significant. This provides reassuring evidence that persistence of participation cannot explain our result.

Another possibility is that our estimates capture coincidental aggregate shocks in the Chiefdom. This could be concerning for Shabunda, the Raia Mutomboki recruited in large numbers at the end of the period. In column (3), we estimate equation 1 including as controls

indicators for Chiefdom interacted with years (the district of Shabunda has two Chiefdoms, and only one is in our sample). The coefficient remains similar, and p-value is 0.10.

A remaining possibility is that attacks by foreign armed groups on the household are simply a proxy for war in the village. Column (4) includes an indicator for whether a territorial conquest occurs in the village as control. If foreign armed group attacks on household members simply capture the effect of war through other channels, the coefficient on foreign attack should vanish. The coefficient is slightly larger, and remains statistically significant. Thus, recrudescence of violence due to war cannot explain our results.

In sum, we find no evidence that attack endogeneity, measurement error, migration, and other leading confounds can explain our main result. With supporting evidence that the coefficient captures a causal relationship from household victimization to respondent participation, we now turn to establishing the channels that may explain it.

7 Examining causal channel: altruistic preferences

We now turn to examining our proposed causal channel for our finding that foreign armed group attack on the household leads to greater propensity to participate in militia.

7.1 Formation of preferences to join a Congolese militia

In this section, we examine whether the baseline estimate reflects preference formation.

Our analysis first examines whether the effect arises from the events that are most traumatic. We do so by examining the actions perpetrators take during the attack, singling out those that are most gruesome. We isolate whether the attack is especially traumatic by whether the attackers targeted the respondent's spouse, whether, they perpetrated sexual violence against a member of the household and, if they did, whether the victim was the spouse, and whether the attackers expropriated the household. Estimates of equation 1 using foreign armed group attacks with these characteristics are reported in Table IV, Panel A.

Column (1) is the benchmark. Column (2) shows the effect for foreign armed group attacks that victimized the respondent’s spouse. The coefficient is almost twice as large, and significant at the 1% level. Table notes also indicate that it is significantly larger than the baseline estimate (“P-value: severe vs. not severe”). Given the control mean is 2.00, the coefficient implies that foreign armed groups’ attacks with sexual violence on the household lead to a 250% increase in the propensity that the respondent joins a militia in the future. Column (3) shows the effect for foreign armed group attacks in the household in which perpetrators committed sexual violence. The effect is again almost twice as large and significance improves. It is also significantly larger than the baseline estimate. Column (4) then narrows the analysis on foreign armed group attacks in which the perpetrators engaged in sexual violence against the respondent’s spouse.³⁷ The coefficient is 5.04, implying a 242% increase in the propensity to join a Congolese militia, and table notes indicate it is significantly larger than the baseline estimate. Finally, column (5) shows that the effect is also larger if the attack includes theft of household property. In sum, the effect is concentrated in foreign armed group attacks that are especially traumatic.³⁸

However, the analysis in Panel A leaves an open puzzle: if the baseline estimate reflects gruesome attacks by foreign armed groups, would gruesome attacks by other perpetrators have the same result? If our conjectured mechanism holds and the effect reflects parochial altruism, more gruesome attacks by Congolese militia should not translate into a higher propensity to participate in a Congolese militia. We turn to this analysis in Panel B.

Columns (1)–(5) of Panel B replicate the analysis of Panel A, but use attacks perpetrated on the household by Congolese militia as attack indicator. The columns replicate the categories from Panel A. We find that none of the attacks has a significant effect on the propensity to join a Congolese militia, and some effects are even negative. This allows us to confirm that the baseline estimate reflects the formation of preferences to fight against foreign armed groups, forged through foreign armed groups’ gruesome attacks.

³⁷Reports of sexual violence on children are less than 10% of the 106 attacks on children.

³⁸Figure D.6 in the online appendix shows that the first attack has the largest effect.

We then analyze the role of the respondent’s age at the time of victimization. Traumatic events are more likely to forge the respondent’s preferences when he is young. Thus, to provide supporting evidence that our estimate reflects the formation of preferences, we analyze whether the effect arises from attacks that take place at an early age. Figure III presents the estimates of equation 1, decomposing the foreign armed group attacks by cohort. Specifically, we estimate the coefficient on I_{it}^{Victim} at different 5-year age intervals of the respondent when the attack takes place. To do so, we interact I_{it}^{Victim} with different indicators for whether the attack took place when the age of the respondent belonged to a given age interval. We find that the effect is the largest (and the only one statistically significant) in the 20–24 age range, and the magnitude of the effect monotonically decreases for older ages, even as the sample size increases. Because the Congolese wars only begin in 1996, victimization at ages below age 20 is extremely infrequent in the sample (it only affects .7% of observations). Thus, the effect of victimization at age below 20 is imprecisely estimated. Household victimization at younger ages (20–24), has a larger impact than at 25–29, which has a larger impact than at 30–34.³⁹ This pattern is consistent with the formation of preferences.

Finally, our analysis examines effect stability. If our baseline estimate reflects the formation of preferences, then the effect should persist over time. Figure IV presents the event study analysis. The figure reports the coefficients for the leads and lags of the attack indicator, using as baseline reference the year before the foreign armed group attack. The brackets indicate confidence intervals allowing to conclude whether the propensity to participate is different than in the year preceding the attack. We find that the effect persists more than five years after the household victimization, consistent with the formation of preferences.

In sum, this section has shown that the effect of foreign armed group attacks on the household is consistent with the formation of stable preferences to join a Congolese militia, formed at young ages. In what follows, we examine their parochially altruistic origins.

³⁹Figure D.7 implements two extensions. In Panel A, we show the results only for respondents who were 18 or older during the Second Congo War, and thus who were at an age where they could participate when militia began forming. Panel B uses only data from the second attack experienced. The effects are similar.

7.2 Parochial altruism: Household, neighbors, and community

We now examine whether such preferences reflect parochial altruism. In this setting, households' reference group is often the community. Thus, if the effect reflects parochial altruism, then attacks that also harm the community's core should have a particularly strong effect. In eastern DRC, customary chiefs play an important role in village life, and are often considered as legitimate representatives of the community (Hoffmann et al., 2020). An attack on the village chief is thus often perceived as an attack on the community's core.

Table V presents the estimates from equation 1, decomposed by whether the village chief is victimized during the household victimization event perpetrated by foreign armed groups, whether other households in the same village, or in the same Chiefdom, are also victimized. Column (1) presents the baseline estimate as benchmark. Column (2) estimates equation 1, including as control variable an indicator for whether the village chief is attacked.⁴⁰ We find that the effect of chief victimization is similar to that of household victimization, yet the baseline estimate remains large and significant.⁴¹ Column (3) includes, in addition, an indicator for whether other households in the village are also attacked. We find that the effect of other households also being attacked is 40% smaller than an attack on his household, and that both the baseline estimate and the coefficient on other households are statistically significant. Column (4) also includes whether other villages in the same Chiefdom are attacked. We find that attacks on other villages have no significant effect. Thus, the effect of a foreign armed group attack is monotonically decreasing if it targets, respectively: the village chief or household members, other villagers, other villagers of the Chiefdom. This is consistent with parochial altruism towards family and community.

In sum, the baseline estimate is consistent with the interpretation that household victimization forges preferences to join anti-foreign armed group militia, and that these reflect altruistic concern for individuals closest to the respondent, and community symbols.

⁴⁰Village chief attack includes all civilian attacks by foreign groups in which the chief is victimized, not confined to attacks that victimize the respondent's household.

⁴¹60% of foreign armed group attacks on the household also victimized the village chief (206 vs. 127 did).

8 Assessing the merit of alternative causal channels

We now examine the merit of remaining alternative causal channels. Section E presents a simple household dynamic optimization model to guide our analysis of causal channels.

The model developed in Section E illustrates two important alternative channels through which armed group attacks on the household may influence participation. First, the attacks may lead to a destruction of household assets, especially if there is theft, and a lower stock of assets can influence future occupational choices. Specifically, with a lower stock of assets, the returns to income and to investment are larger, hence the incentives to participate in occupations that generate income, even if those are risky, are stronger. If armed groups provide high income, this channel could explain our results.⁴² Second, the attacks may influence the respondents' attitudes towards attacks by making them more sensitive about risk of attacks. If this channel operates, respondents from victimized households may be more likely feel insecure, which could lead them to join a militia to obtain higher security.

In addition, attacks on other household members may be correlated with attacks on the respondent himself. This could be an alternative causal channel, since attacks on the respondent likely affect the respondent's capabilities, directly influencing his choice to join armed groups independently of their effect on preferences (for instance, through handicap).

Table VI examines the merits of these alternative causal paths. Column (1) is the benchmark. Column (3) includes household investment in year t as a control. Since investment is only available for South Kivu, Column (2) shows the benchmark in South Kivu. Column (4) also controls for whether the respondent overreports attacks. Column (5) estimates equation 1 controlling for whether the respondent is physically harmed. Column (6) includes all these controls. The baseline estimate and significance are preserved across specifications.

Thus, we find no merit for the possibility that foreign armed group attacks on the household lead respondents to participate in a Congolese militia *because* they affect their incentives to invest, attitudes toward risk, or the respondent's ability to participate in the labor market.

⁴²This is unlikely, since we have shown that the effect is concentrated on militia.

9 External validity: the effect in weak and strong states

We now analyze the role of the state in mediating our baseline estimate. The state, represented militarily by the state forces, typically aims to maintain a monopoly of violence, and therefore repress alternative organizations of violence, including Congolese militia. Our baseline estimate may thus be positive because, on average, the state is relatively weak and cannot repress Congolese militia to form. This has implications for external validity: if our result reflects state weakness, then it may not apply in well-functioning states.

In our study period, there are two episodes in which the state forces withdraw, which we henceforth call weak state episodes (see Section 2).

Figure V shows the trends in presence of state forces and of fraction of individuals who join a militia in the period. We define state forces as the Congolese army or the Rassemblement Congolais pour la Democratie (RCD), a group who took over the apparatus of the state and acted as the state forces during all of the Second Congo War, 1998–2004. Panel A shows the trends for sample of all villages. Since the start of the Congo wars in 1996, there are two episodes of state weakness, in which the fraction of villages with a state force drops: 2003–5 and 2011–13.⁴³ Participation is more pronounced in weak state episodes. As Section 2 suggested that state weakness episodes affected more the district of Shabunda, Panel B shows that the state weakness episodes in Shabunda are more pronounced.

The times-series in Figure V suggest that in years with strong state, participation is likely to be costlier (i.e., repressed by the state forces).⁴⁴ We now analyze this possibility formally.

To separately identify the effect of past attacks in weak and strong state episodes, we estimate equation 1, and include, in addition to the foreign armed group attack indicator,

⁴³These correspond to the security vacuums previously identified—notably the integration of Mai-Mai armed groups into the National Army, and the regimentation process of the national army. Before the wars, the state also had a weak presence. However, even if that made participation easier, the environment was also less militarized at that time and thus the state weakness episode is less relevant.

⁴⁴In fact, historical evidence shows how the state forces, both the Congolese army and the RCD, engaged in drastic counterinsurgency campaigns aimed at eradicating Congolese militia in the areas that they controlled (Stearns, 2011). While the RCD was extremely violent in repressing militia, the Congolese army continues to fight many Mai-Mai's groups today to assert their territory.

I_{it}^{Victim} , its interaction with an indicator taking value 1 if no state forces are present in the village in that year (“weak state”), as well as the indicator alone. Table VII presents the estimates. The dependent variable in columns (1)–(3) is an indicator taking value 1 if the respondent participates in a Congolese militia. Column (1) is the benchmark.

Column (2) implements this fully saturated model. We find that the effect of an attack by a foreign armed group arises *exclusively* through years in which there is no state force present. Indeed, the coefficient on I_{it}^{Victim} is negative and insignificant, indicating that there is no effect on joining militia in years in which state forces are present. The coefficient on the interaction term, 7.12 pp., is larger than the baseline estimate, indicating that foreign armed group attacks cause respondents to join militia only in years in which state forces are absent. This implies that a foreign armed group attack on the household causes a 766% increase in the propensity that the respondent joins a militia in a weak state year.

Columns (3) and (4) exploit plausibly exogenous variation in the weakness of the state to confirm that the the effect of state presence/absence on our baseline estimate is causal. We use information on dates and locations of military redeployment episodes in Shabunda, which created state weakness.⁴⁵ Column (3) is the reduced form. It replicates the specification of column (2), but instead of using the indicator “weak state,” it uses an indicator for whether the observation is in Shabunda in the years 2003–05 and 2011–13. We find similar results. Column (4) instruments for whether the state forces are present using an indicator for whether the observation is in Shabunda in the years 2003-2005 and 2011-2013 as an instrument. The first stage is strong (F-stat 19.51), and the main coefficient (Past attack * weak state) is large and significant. The conclusions are preserved.

In sum, participation is the byproduct of the interaction between preferences to fight foreign armed groups, forged past victimizations, and the weakness of the state.

⁴⁵In particular, we code as 1 individual-year observations in which the respondent lives in Shabunda in the 2003–05 episode in which the “state force” RCD vacated Shabunda drastically as a result of the Sun City peace agreement, or in the 2011–13 episode in which the regimentation process led the Congolese army to vacate Shabunda. These shocks, mandated by arguably exogenous military strategy reasons that were unrelated to the specific location in which they occurred, are potentially valid instruments for state weakness.

10 Preferences, institutions, or income?

In previous sections, we have shown that foreign armed group attacks on household members lead to the formation of intrinsic motivation to join a Congolese militia and fight foreign armed groups. In this section, we benchmark the magnitude of our effect to that of the effect of state weakness, and to that of the effect of changes in income outside armed groups.

10.1 Preferences to fight vs. state weakness

We now compare the independent role of the state and that of preferences in explaining the decision to participate in a Congolese militia. Table VII decomposed the main result into episodes of state presence and state absence in the villages. The same specification allows to measure the independent effect of state absence on participation in a Congolese militia, and compare it to the effect of preferences formed through household victimization.

The coefficient in Column (2) on the indicator “weak state” measures the effect the absence of state forces on whether the respondent participates in a Congolese militia. The coefficient, 3.82, is large and significant at the 1% level. Thus, weak state episodes are associated with a higher propensity to participate in a Congolese militia. The magnitude suggests that weak state or household victimization have a similar effect on participation.

Furthermore, the coefficient on “Past attack * Weak state” is 7.12, thus the effect of state absence is even larger if the preferences for rebellion are already seeded in the population. It implies that, on average, absence of the state leads to an increase in the propensity that individuals participate in a Congolese militia by between 3.82 pp. (if they have not been victimized) and 10.94 pp. (if they have). Columns (3) and (4) reinforce this conclusion.

In sum, past household victimization by foreign armed groups, and state absence, lead to a comparable increase in the propensity that respondents participate in a Congolese militia. Specifically, the effect of withdrawing state forces on whether respondents participate in a Congolese militia (without victimization) is 149% that of victimization (without withdrawal).

10.2 Preferences vs. economic incentives: income fluctuations

In this section, we benchmark the magnitude of the effect of a foreign armed group attack on the household to the effect of income fluctuations outside armed groups.

Consider the trade-off of an individual deciding whether to leave the productive sector to enter an armed group. If the income he can obtain outside the armed group is large, while the material or nonmaterial benefits he derives from joining remain intact, the decision to join the armed group entails higher material losses, the opportunity cost of income he could otherwise obtain. Thus, if economic incentives play any role at influencing this decision, higher income outside the armed groups would reduce the propensity of individuals who join armed groups. This is the standard framework in economics for analyzing the decision to enter armed groups (Dube and Vargas, 2013). We use this as a benchmark.

To measure the effect of non-armed group income on the decision to join armed groups, we exploit the presence of large variations in the world price of minerals, for which deposits are found in some of the villages of the sample and for which DRC is a price taker. Gold is easy for workers to conceal, making it often difficult for armed groups to tax it. As a result, in gold mining villages, an increase in the world price of gold translates almost 1-to-1 to an increase in the price received by workers in mining, but changes little the revenues of armed groups in gold villages.⁴⁶ We take advantage of this feature, and use a large change in the price of gold in this period, to quantify the effect of non-armed group income on the propensity to join an armed group.⁴⁷ We first estimate the reduced form relationship between world prices and propensity to participate in an armed group. We estimate equation 1, but include, in addition, an indicator for whether village j is endowed with mineral m ,

⁴⁶Figure D.8 presents the times-series of the world price of gold. The rise in the world price of gold translates into a rise of the local price of gold, and thus reassures that world price shocks translate into local price shocks with a high rate of passthrough. See Sánchez de la Sierra (2020) for a detailed discussion.

⁴⁷There is also coltan. Coltan is bulky, and thus prone to taxation by armed groups. As a result, an increase in the price of coltan increases both the income of individuals who work in mining in coltan villages, and the income of armed groups—hence, potentially, the income obtained by joining the group. Our analysis confirms this. As a result, the price of coltan does not offer a useful benchmark for economic incentives.

interacted with the world price of mineral m . Specifically, we estimate:

$$Part_{ijt} = \alpha + \gamma I_{it}^{Victim} + \gamma^E I_j^m \times \log(P_t^m) + \alpha_i + \alpha_j + \alpha_t + \alpha_a + \mathbf{X}'_{it}\Gamma + \epsilon_{ijt}, \quad (2)$$

where I_j^m is an indicator taking value 1 for all years if village j has mineral m deposits and P_t^m is the world price of mineral m in year t . As before, this equation includes year, age, and village fixed effects and standard errors are clustered two-ways at the individual and the village-year level. The coefficient on γ^E measures the reduced form relationship between the world price of mineral m and the probability that an individual participates in an armed group. If the world price of mineral m is exogenous and has no spillover effects on villages not endowed with m , γ^E also has a causal interpretation: it measures the effect of a 100% increase in the world price of m on the probability of participation.

The estimates from Equation 2 are presented in Figure VI. Since the results are similar if $I_j^m \times P_t^m$ and I_{it}^{Victim} are included separately, we present the results of the equation including both.⁴⁸ Figure VI presents the coefficient estimates from equation 2 with, as dependent variables, respectively, indicators for whether the individual participates in any armed group, in a Congolese militia, in a foreign armed group, and, since changes in mining sector income may affect the inter-dependent choice of occupations, we also show whether he works mainly in agriculture, mining, government agencies, is in school, or is unemployed.⁴⁹ The reduced form relationship between the price of gold and participation in militia shown in Figure VI indicates that the price of gold, in gold villages, draws individuals away from Congolese militia and into the mining sector. This is the standard opportunity cost channel. The figure also confirms that our baseline estimate survives the inclusion of the price shocks.⁵⁰

⁴⁸Inclusion of $I_j^m \times P_t^m$ for $m = coltan$ does not change the results.

⁴⁹Finally, since the price and experiencing past foreign armed group attacks on the household influence the decision to choose occupations jointly, this figure presents the result using j as main occupation rather than works in occupation j . For South Kivu, we observe what occupations an individual works in each year, and what is his main occupation in that year. For North Kivu, we only observe the latter. In Panel A, we present the result for the main occupation to elicit occupational transitions, since often individuals otherwise work in two occupations. Table C.5, Panel A, replicates the figure in format table and Panel B presents, for robustness, the analysis including indicators for whether the individual works in some occupation.

⁵⁰Table C.5, Panel B confirms this result, and shows that foreign armed group attacks on the household

Table VIII presents the analysis of attacks and price shocks conducted simultaneously. First, we quantify the effect of changes in the local price of gold, using average yearly local price of gold from our survey, and instrumenting it with the world price of gold. We find that it would require a US\$24.1 per gram in the local price of gold to undo the effect of a foreign armed group attack. This is equivalent to 2.9 standard deviations, and an increase in almost 100% of the local price of gold. This implies that one foreign armed group attack on the household requires an increase in US\$24.1 in the local price per gram of gold to be undone, in gold villages. Second, we estimate the equivalent rise in per capita income outside the armed groups that would be necessary to undo the effect of household victimization by foreign armed actors. We use information on the daily production of gold by a gold miner (Geenen, 2013), our data on the taxation of gold activity by local authorities, the GDP per capita of the DRC in that period (we take year 2005) and, assuming miners work 300 days a year, we find that it would take a *permanent* increase in 18.2 times the yearly per capita income to undo the magnitude of the effect of a foreign armed group-perpetrated household victimization. This rise is prohibitive. Its magnitude suggests our baseline estimate presents the formation of “deontological” preferences to fight—those for which there exists no feasible price for which economic incentives can undo them. Kant (2019) referred to these as dignity.

In sum, household victimization by foreign armed groups leads to the formation of preferences for joining Congolese militia. Those only translate into action when the state is weak. To undo it requires a permanent increase in 18.2-fold yearly p.c. income outside militia.

11 Conclusion

This paper provides evidence for the forging of rebels. It illustrates that violent popular movements form as a result of the interaction of intrinsic motivation to take arms and state weakness. The effect of victimization on participation is so large that it would take a prohibitive increase in income to undo it, consistent with formation of deontological preferences.

takes individuals mostly from agriculture and into Congolese militia, suggesting it is a peasant phenomenon.

This is a story of community and family victimized by foreign armed groups. Violations perpetrated by foreign armed groups have generated among the relatives of the victims a desire and possibly a moral conviction to fight back. It is also a story of peasants who choose to fight against those who exploited the absence of their protection. It shows the forging of rebels through the forging of preferences, long studied in novels and in historical accounts, but for which there is to date no evidence in the field of economics. It shows that nonmaterial motives can explain conflict, a high-stake developmental outcome. We hope that this settles the debate in economics about the role of nonmaterial motivations for conflict.

Future research should explore the mechanism in more detail. We provided evidence supporting the formation of intrinsic motivation, but what type of intrinsic motivation is likely to vary across cultural context, and in implications. For instance, our results are consistent with the violation of dignity, with honor, but also with revenge (Balcells, 2017), retaliation, the formation of concern for social image to be seen fighting. The study of some of those preferences remains outside the scope of economic analysis, despite their importance in various other areas of social science inquiry. Economic analysis would greatly benefit from understanding those emotions and how they relate to preferences, beliefs, and choices.

Our study also leaves an important aspect unstudied. We have chosen to analyze the decision to join armed groups as reflecting individual decision-making. However, in reality, the processes of militia mobilization are largely social. They involve local institutions, social networks, and the activation of power and often subtle forms of within-community coercion. Future research would greatly benefit from studying how these intrinsic motivations interact with the processes of mobilization. We leave those questions for future research.

References

Arjona, Ana, “Local Orders in Warring Times: Armed Groups’ and Civilians’ Strategies

- in Civil War,” *Qualitative Methods*, Spring 2008.
- Ashraf, N., O. Bandiera, and S.S Lee**, *Do-Gooders and Go-Getters: Career Incentives, Selection, and Performance in Public Service Delivery* Economic organisation and public policy discussion papers, STICERD, LSE, 2014.
- Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack**, “No Margin, No Mission? A field Experiment on Incentives for Public Service Delivery,” *Journal of Public Economics*, 2014, *120* (C), 1–17.
- Atran, Scott**, “The Devoted Actor: Unconditional Commitment and Intractable Conflict across Cultures,” *Current Anthropology*, 6 2016, *57* (S13), S192–S203.
- Balcells, Laia**, “The Consequences of Victimization on Political Identities: Evidence from Spain,” *Politics & Society*, 2012, *40* (3), 311–347.
- , *Rivalry and Revenge The Politics of Violence during Civil War*, Cambridge: Cambridge University Press, 4 2017.
- Barea, Arturo**, *The Forging of a Rebel*, London : Fontana, 1984.
- Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts**, “Can War Foster Cooperation?,” *Journal of Economic Perspectives*, September 2016, *30* (3), 249–74.
- Blattman, Christopher and Edward Miguel**, “Civil War,” *Journal of Economic Literature*, March 2010, *48* (1), 3–57.
- Brabant, Justine**, “*Qu’on nous laisse combattre, et la guerre finira*”: Avec les combattants du Kivu, Paris: La Découverte, 2016.
- Bénabou, Roland and Jean Tirole**, “Intrinsic and Extrinsic Motivation,” *The Review of Economic Studies*, 2003, *70* (3), 489–520.

- Bó, Ernesto Dal, Frederico Finan, and Martín A. Rossi**, “Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service,” *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.
- , – , **Olle Folke, Torsten Persson, and Johanna Rickne**, “Who Becomes A Politician?*,” *The Quarterly Journal of Economics*, 06 2017, *132* (4), 1877–1914.
- Collier, Paul**, “On the Economic Consequences of Civil War,” *Oxford Economic Papers*, 1999, *51* (1), 168–183.
- **and Anke Hoeffler**, “On Economic Causes of Civil War,” *Oxford Economic Papers*, 1998, *50* (4), 563–573.
- **and** – , “Greed and grievance in civil war,” *Oxford Economic Papers*, 2004, *56* (4), 563–595.
- Deserranno, Erika**, “Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda,” *American Economic Journal: Applied Economics*, 01 2019, *11*, 277–317.
- Dube, Oeindrila and Juan Vargas**, “Commodity Price Shocks and Civil Conflict Evidence From Colombia,” *Review of Economic Studies*, 2013.
- Fearon, James D.**, “Rationalist Explanations for War,” *International Organization*, 1995, *49* (3), 379–414.
- Geenen, Sara**, “Who Seeks, Finds: How Artisanal Miners and Traders Benefit from Gold in the Eastern Democratic Republic of Congo,” *European Journal of Development Research*, June 2013, *25*, 197 – 212.
- Gould, Roger V.**, “Collective Action and Network Structure,” *American Sociological Review*, 1993, *58* (2), 182–96.

Gurr, Ted Robert, *Why Men Rebel* 1970.

Heineman, Matthew, “Cartel Land, Documentary,” 2015.

Hoeffler, Anke, “What are the costs of violence?,” *Politics, Philosophy & Economics*, 2017, 16 (4), 422–445.

Hoffmann, Kasper, “Myths Set in Motion: The Moral Economy of Mai Mai Governance,” in Ana Arjona, Nelson Kasfir, and Zachariah Mampilly, eds., *Rebel Governance in Civil War*, Cambridge: Cambridge University Press, 2015.

– **and Judith Verweijen**, “Rebel rule: A Governmentality Perspective,” *African Affairs*, 09 2018, 118 (471), 352–374.

– **and Koen Vlassenroot**, “Armed Groups and the Exercise of Public Authority: the Cases of the Mayi-Mayi and Raya Mutomboki in Kalehe, South Kivu,” *Peacebuilding*, 2014, 2 (2), 202–220.

– , – , **and Emery Mudinga**, “Courses au Pouvoir: the Struggle Over Customary Capital in the Eastern Democratic Republic of Congo,” *Journal of Eastern African Studies*, 1 2020, 14 (1), 125–144.

– , – , **and Gauthier Marchais**, “Taxation, Stateness and Armed groups : Public Authority and Resource Extraction in Eastern Congo,” in Christian Lund and Michael Eilenberg, eds., *Rule and Rupture : State Formation Through the Production of Property and Citizenship*, Wiley-Blackwell, 2017, pp. 235–256.

Humphreys, Macartan and Jeremy M. Weinstein, “Who Fights? The Determinants of Participation in Civil War,” *American Journal of Political Science*, 4 2008, 52 (2), 436–455.

- Jackson, Matthew O. and Massimo Morelli**, “The Reasons for Wars: An Updated Survey,” in Christopher J. Coyne and Rachel L. Mathers, eds., *The Handbook on the Political Economy of War*, Chapters, Edward Elgar Publishing, 2011, chapter 3.
- Jackson, Stephen**, “Sons of which soil? The Language and Politics of Autochthony in Eastern DR Congo,” *African studies review*, 2006, 49 (2), 95–124.
- Johnson, Dominic, Simone Schlindwein, and Bianca Schmolze**, *Les FDLR, histoire d’une milice rwandaise : des forêts du Kivu aux tribunaux de l’Allemagne*, Berlin: Christoph Links Verlag, 2016.
- Jourdan, Luca**, “Being at War, Being Young: Violence and Youth in North Kivu,” in Koen Vlassenroot and Timothy Raeymaekers, eds., *Conflict and Social Transformation in Eastern DR Congo*, Gent, Belgium: Academia Press, 2004.
- , “Mayi-Mayi: Young Rebels in Kivu, DRC,” *Africa Development*, 2011, XXXVI (3-4), 89–111.
- Kalyvas, Stathis N.**, *The Logic of Violence in Civil War* Cambridge Studies in Comparative Politics, Cambridge University Press, 2006.
- **and Matthew Adam Kocher**, “How “Free” is Free Riding in Civil Wars?: Violence, Insurgency, and the Collective Action Problem,” *World Politics*, 1 2007, 59 (2), 177–216.
- Kant, Immanuel**, *Groundwork for the Metaphysics of Morals*, Oxford University Press, 2019.
- Karing, Anne**, “Social Signaling and Childhood Immunization,” *Working paper*, 2021.
- Khan, Muhammad Yasir**, “Mission Motivation and Public Sector Performance: Experimental Evidence from Pakistan,” *Job Market Paper*, 2021.
- Levitt, Steven and Sudhir Alladi Venkatesh**, “An Economic Analysis of a Drug-Selling Gang’s Finances,” *The Quarterly Journal of Economics*, 2000, 115 (3), 755–789.

- Lowes, Sara and Eduardo Montero**, “Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State,” with E. Montero,” *Quarterly Journal of Economics*, 2021, *forthcoming*.
- Marchais, Gauthier**, “He who touches the weapon becomes other: a study of participation in armed groups in South Kivu, Democratic Republic of the Congo,” *PhD thesis*, 2016.
- Parkinson, Sarah Elizabeth**, “Organizing Rebellion: Rethinking High-Risk Mobilization and Social Networks in War,” *American Political Science Review*, 2013, 107 (3), 418–432.
- Petersen, Roger Dale**, *Resistance and Rebellion: Lessons from Eastern Europe*, New York: Cambridge University Press, 2001.
- Popkin, Samuel L.**, *The rational peasant : the Political Economy of Rural Society in Vietnam*, University of California Press, 1979.
- Sánchez de la Sierra, Raúl**, “On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo,” *Journal of Political Economy*, 2020, 128 (1), 32–74.
- , “Whither formal contracts?,” *Econometrica*, 2021, *forthcoming*.
- Sanín, Francisco Gutiérrez and Elisabeth Jean Wood**, “Ideology in Civil War: Instrumental Adoption and Beyond,” *Journal of Peace Research*, 2014, 51 (2), 213–226.
- Scacco, Alexandra**, *Anatomy of a Riot Participation in Ethnic Violence in Nigeria* 2016.
- Schubiger, Livia Isabella and Matthew Zelina**, “Ideology in Armed Groups,” *PS - Political Science and Politics*, 10 2017, 50 (4), 948–951.
- Staniland, Paul**, *Networks of rebellion : explaining insurgent cohesion and collapse*, Ithaca and London: Cornell University Press, 2014.
- Stearns, Jason**, *Dancing in the Glory of Monsters*, New York, NY: Public Affairs, 2011.

– **and Raphael Botiveau**, “Repenser la crise au Kivu : mobilisation armée et logique du gouvernement de transition,” *Politique Africaine*, 2013, 1 (129), 204.

Stearns, Jason K., “From CNDP to M23: The evolution of an armed movement in Eastern Congo,” Technical Report, Report by the Usalama Project: Understanding Armed Groups. Rift Valley Institute, London 2012.

Stearns, Jason K., “North Kivu: The background to conflict in North Kivu province of eastern Congo,” Technical Report, Report by the Usalama Project: Understanding Congolese Armed Groups. Rift Valley Institute, London 2012.

Stearns, Jason K., “Raia Mutomboki: The Flawed Peace Process and the Birth of an Armed Franchise,” Technical Report, Report of the Usalama Project: Understanding Congolese Armed Groups. The Rift Valley Institute, London 2013.

Stearns, Jason K., Judith Verweijen, and Maria Eriksson Baaz, “The National Army and Armed Groups in the Eastern Congo: Untangling the Gordian Knot of Insecurity,” Technical Report, Report by the Usalama Project: Understanding Congolese Armed Groups. Rift Valley Institute, London 2013.

Stys, Patrycja, “FDLR: Past, Present, and Policies,” Technical Report, Social Science Research Council, New York 2014.

Verweijen, Judith, “A Microcosm of Militarization: Conflict, governance and armed mobilization in Uvira, South Kivu,” Technical Report, Report by the Usalama Project II: Governance in Conflict. Rift Valley Institute, London 2016.

Viterna, Jocelyn, *Women in War: The Micro-processes of Mobilization in El Salvador*, Oxford University Press, 11 2013.

Viterna, Jocelyn S., “Pulled, pushed, and persuaded: Explaining women’s mobilization into the Salvadoran guerrilla army,” *American Journal of Sociology*, 7 2006, 112 (1), 1–45.

- Vlassenroot, Koen**, “Citizenship, Identity Formation Conflict in South Kivu: The Case of the Banyamulenge,” *Review of African Political Economy*, 2002, 29 (93/94), 499–515.
- , “South Kivu: identity, territory, and power in the eastern Congo,” Technical Report, Report by the Usalama Project: Understanding Congolese Armed Groups. Rift Valley Institute, London 2013.
- **and Frank Van Acker**, “War as Exit from Exclusion? The Formation of Mayi-Mayi Militias in Eastern Congo,” *Afrika Focus*, 8 2001, 17 (1-2), 51–77.
- Vogel, Christoph**, “Contested Statehood, Security Dilemmas and Militia Politics: the Rise and Transformation of Raia Mutomboki in Eastern DRC,” in Filip Reyntjens, Steph Vandeginste, and Marijke Verpoorten, eds., *L’Afrique des Grands Lacs. Annuaire.*, Paris: L’Harmattan, 2014.
- **and Jason K Stearns**, “Kivu’s Intractable Security Conundrum, Revisited,” *African Affairs*, 08 2018, 117 (469), 695–707.
- , **Giovanni Salvaggio, Pierre Boisselet, and Jason K Stearns**, “The Landscape of Armed Groups in Eastern Congo: Missed Opportunities, P-rotracted Insecurity and Self-Fulfilling prophecies,” Technical Report, Kivu Security Tracker - Congo Research Group - NYU Center on International Cooperation, New York 2021.
- Weinstein, Jeremy M.**, *Inside Rebellion: The Politics of Insurgent Violence*, Cambridge, UK: Cambridge University Press, 2007.
- Wood, Elisabeth Jean**, *Insurgent Collective Action and Civil War in El Salvador*, Cambridge University Press, 2003.
- Wood, Elizabeth**, “Variation in Sexual Violence during War,” *Politics and Society*, September 2006, 34 (3), 1–21.

Tables

Table I: Characteristics of Non-rebels and Rebels

	Not Part. (1)	Start Part. (2)
Obs. (Age ≥ 15 , 1995 $\leq t \leq 2013$)	24,307	296
<i>Panel A: Participation characteristics per participation episode</i>		
Participates in Congolese militia in year t (%)		78.38
Participates in foreign armed group in year t (%)		11.15
Participates in Congolese army in year t (%)		2.03
Average duration		2.34
<i>Panel B: Socio-demographic background per individual-year obs</i>		
Age in year t	32.14	31.28***
Married in year t (%)	44.79	18.92***
Works primarily in mining in year $t - 1$ (%)	14.01	16.27
Works primarily in agriculture in year $t - 1$ (%)	46.81	44.84
Works primarily as civil servant in year $t - 1$ (%)	6.70	11.51
Works primarily in school in year $t - 1$ (%)	9.84	6.75
Unemployed in year $t - 1$ (%)	22.64	20.63
<i>Panel C: Economic status per individual-year obs</i>		
Wealth at birth (z-score, only $t = 2012$)	-0.02	0.05
Asset stock in year $t - 1$ (z-score)	-0.02	-0.20
Investment in year $t + 1$ (z-score)	0.09	0.38
<i>Panel D: Attack history per individual-year obs</i>		
Experienced attack on own HH by foreign armed group before year t (%)	8.91	22.30*
Experienced attack on own HH by Congolese militia before year t (%)	2.92	2.36
Experienced attack on own HH by Congolese army before year t (%)	0.17	0.34

Notes: This table shows descriptive statistics for observations where respondents start participating in any armed group versus those where respondents do not participate in year t . We exclude from comparison 488 observations where participants stay in the armed group after the first year. Economic indices are only computed in South Kivu, where the data are available. We indicate the difference between Column 1 and 2 (P-value: *** 0.01, ** 0.05, * 0.10), computed after including village FE and year FE, and clustered two-way at the individual respondent and the village*year level. Construction of economic indices is made using principal component analysis on following variables. For the wealth at birth index, we use stock of cows at birth, stock of goats at birth, stock of pigs at birth, stock of lands at birth, relation to village chief, number of father's wives. For the asset stock index, we use stock of cows, stock of goats, stock of pigs, stock of lands. For the investment index, we use purchase of cows, purchase of goats, purchase of pigs, purchase of lands.

Table II: Baseline Estimate

Past attack on HH	Recruiter groups and identities		
	Congolese militia (1)	Foreign armed group (2)	Any armed group (3)
By foreign armed group	2.55** (1.04) [0.014]	-0.46 (0.45) [0.430]	2.51** (1.21) [0.028]
By Congolese militia	0.69 (1.32) [0.615]	-0.30 (0.41) [0.485]	0.97 (1.67) [0.609]
By any armed groups	1.78** (0.87) [0.027]	-0.38 (0.39) [0.384]	1.73* (1.01) [0.100]
Control mean	1.88	0.33	2.72
Obs.	25,060	25,060	25,060

Notes: This table presents OLS estimates of equation 1 for different perpetrators (by foreign armed groups, by Congolese militia, and by any armed group respectively) and for different recruiters (Congolese militia, foreign armed group, and any armed group recruiter). The dependent variable in columns (1)–(3) is an indicator for whether the respondent joins in a given year a foreign armed group, a Congolese militia, or any armed group, respectively. The explanatory variable in each line is an indicator for whether the respondent’s household has been attacked by foreign armed groups, a Congolese militia, or any armed group, respectively. Each coefficient is estimated from a separate regression. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household from any armed group before year t . P-value: *** 0.01, ** 0.05, * 0.10. Figure D.5 shows the graphical representation of the coefficients from Column (1)—that is, the effect of past attack on the household by different perpetrators on participation in a Congolese militia.

Table III: Baseline Estimate—Examining Potential Confounders

	Participation in Congolese militia in year t (%)			
	(1)	(2)	(3)	(4)
<i>Panel A1. Endogeneity of armed group attack, controlling confounding factors</i>				
Past attack on HH by foreign armed group	2.55** (1.01)	4.66** (2.24)	5.25*** (1.92) 0.33 (0.23)	4.79*** (1.76)
Asset stock $t - 1$ (z-score)				-72.08*** (14.80)
Predicted attack likelihood				
Attack type	All	Intention: Not pillage	All	All
Control mean	1.89	2.19	2.95	2.81
Obs.	25,060	25,060	15,032	16,312
<i>Panel A2. Endogeneity of armed group attack, instrumental variable approach</i>				
Past attack on HH by foreign armed group	2.55** (1.01)	323.60** (164.63)	26.61* (15.45)	30.19* (17.44)
IV	Main	Lived in Shabunda between 1998–2002 or 2010–11	Lived in village controlled by foreign armed group between 1998–2002 or 2010–11	× From rich family
F-stat		3.96	19.79	8.34
Control mean	1.89	2.57	2.51	2.55
Obs.	25,060	18,135	18,337	17,909

Notes: This table estimates equation 1 using different specifications to examine potential confounders of main specification 1, focused on attacks on household by foreign armed group and participation in a Congolese militia. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household from foreign armed group before year t , controlling other variables at mean values. P-value: *** 0.01, ** 0.05, * 0.10.

a. Panel A1: Column (2) uses only attacks without intention to pillage. Column (3) controls for asset stock index in year $t - 1$. Column (4) controls for predicted attack likelihood using unbalanced variables from Table C.2.

b. Panel A2: Column (2) uses as instrument whether respondent lived in Shabunda during Second Congo war (1998–2002) or during Kimia II operation (2010–11). Column (3) uses as instrument whether respondent lived in a village controlled by foreign armed group during Second Congo war or Kimia II operation. In column (4), we use the same instrument as in Column (3) but interact it with an indicator for whether the household originates from a rich family.

	(1)	(2)	(3)	(4)	(5)
<i>Panel B. Measurement error</i>					
Past attack on HH by foreign armed group	2.55** (1.01)	3.29* (1.77)	2.96* (1.70)	5.40*** (2.08)	36.84** (17.08)
Control mean	1.89	2.06	1.53	2.89	2.52
Obs.	25,060	25,060	17,401	15,821	17,358
Robustness test	Main	Cross-validated attacks	Weighted by attack precision	Weighted by working memory	IV: past village attack by foreign group
F-stat					13.09
<i>Panel C. Selection bias of migration</i>					
Past attack on HH by foreign armed group	2.55** (1.01)	4.02*** (1.48)	4.12*** (1.44)		
Control mean	1.89	2.28	1.89		
Obs.	25,060	12,710	25,060		
Robustness test	Main	Villagers who never leave the sample	Control out-of-sample migration		
<i>Panel D. Other threats to inference</i>					
Past attack on HH by foreign armed group	2.55** (1.01)	1.04** (0.47)	2.14 (1.31)	2.68*** (1.01)	
Control mean	1.89	0.83	1.89	2.16	
Obs.	25,060	25,060	17,886	25,060	
Robustness test	Main	Persisting participation	Control Chiefdom-year trend	Control conquest	

c. Panel B: Column (2) uses only cross-validated attacks. Column (3) weights observations by precision of attack reporting. Column (4) weights observations by working memory test. Column (5) uses village-level information on attacks by foreign armed group as instrument.

d. Panel C: Column (2) uses only observations where respondent stayed in the sample of interview villages. Column (3) uses the full sample but controls for whether the respondents lived outside the interview villages.

e. Panel D: Column (2) uses only the start of each participation episode as dependent variable. Column (3) controls for Chiefdom-year trend. Column (4) controls for territorial conquest.

Table IV: Formation of Preferences: Traumatic Events—Gruesome Attacks

	Participation in Congolese militia in year t (%)				
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Attack by foreign armed group</i>					
Past attack on HH by	2.55**	4.54***	4.25***	5.04***	2.98***
foreign armed group $\times S$	(1.01)	(1.30)	(1.50)	(1.80)	(1.10)
Attack severity S :					
General attack	✓	✓			
Sexual violence			✓	✓	
On wife		✓		✓	
Theft					✓
P-value: severe vs. not severe		0.00	0.07	0.05	0.00
Control mean	1.89	2.00	2.02	2.08	1.87
Obs.	25,060	25,060	25,060	25,060	25,060
<i>Panel B: Attack by Congolese militia</i>					
Past attack on HH by	0.52	0.62	-0.39	-0.05	0.92
Congolese militia $\times S$	(1.66)	(1.22)	(0.71)	(0.85)	(1.28)
Attack severity S :					
General attack	✓	✓			
Sexual violence			✓	✓	
On wife		✓		✓	
Theft					✓
P-value: severe vs. not severe		0.91	0.38	0.57	0.05
Control mean	2.20	2.21	2.27	2.26	2.20
Obs.	25,060	25,060	25,060	25,060	25,060

Notes: This table presents OLS estimates of equation 1 for attacks of different severity. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups (Panel A) and by Congolese militia (Panel B), respectively. We include observations between 1995 and 2013 above age 15 at year t . Column (2) focuses on attacks where respondent's spouse was attacked. Column (3) focuses on attacks where perpetrators committed sexual violence in the household. Column (4) focuses on attacks where sexual violence was against respondent's spouse. Column (5) focuses on attacks involving theft of household property. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t . P-value: *** 0.01, ** 0.05, * 0.10.

Table V: Parochial Altruism as Origin—Victimization of Household and Community

Past attack by foreign armed group	Participation in Congolese militia t (%)			
	(1)	(2)	(3)	(4)
On own household	2.55** (1.01)	1.84* (0.96)	1.75* (0.95)	1.73* (0.95)
On village chief		2.62*** (0.76)	2.44*** (0.75)	2.45*** (0.75)
On other households, same village			1.17* (0.66)	1.14* (0.67)
On other villages, same Chiefdom				0.37 (0.65)
Control mean	1.89	1.21	0.92	0.76
Obs.	25,060	25,060	25,060	25,060

Notes: This table presents OLS estimates of equation 1 for attacks in different scopes of community. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . Column (2) controls for attacks where the village chief was attacked. Column (3) controls for attacks where other respondents in the village were also attacked. Column (4) controls for whether households from other villages in the same Chiefdom are attacked (observations with unknown Chiefdom are grouped under one category). All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t . P-value: *** 0.01, ** 0.05, * 0.10.

Table VI: Examining Alternative Causal Channels

	Participation in Congolese militia in year t (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Past attack on HH by foreign armed group	2.55** (1.01)	4.66*** (1.75)	4.94*** (1.80)	2.58** (1.01)	2.55** (1.01)	5.03*** (1.80)
Investment t (z-score)			0.42** (0.16)			0.42** (0.16)
Overreport any violent event t				0.67 (0.75)		2.31** (1.15)
Past attack on self by foreign armed group, not on other HH members					-0.02 (0.73)	0.35 (1.26)
South Kivu only		✓				
Control mean	1.89	2.81	2.30	1.80	1.81	1.96
Obs	25,060	16,312	14,874	25,060	25,060	14,874

Notes: This table presents OLS estimates of equation 1 for attacks controlling for alternative mechanisms. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . Column (2) shows benchmark results only in South Kivu. Column (3) controls for investment and asset stock in current year t (only coefficient of current investment is shown). Column (4) controls for whether respondent overreports any violent event on village in year t . Overreporting is defined as whether respondent reports a violent event on village in year t , but less than half of the other contemporary villagers observed in the sample report so and there is no corresponding record in village chief survey. Column (5) controls for whether respondent experienced an attack in the past on himself but not on other household members (more discussion in Appendix F). Column (6) controls for all three alternative channels. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t . P-value: *** 0.01, ** 0.05, * 0.10.

Table VII: Formation of Preferences: Effect Stability—Interaction with Weak State Shocks

	Participation in Congolese militia t (%)			
	(1)	(2)	(3)	(4)
Past attack on HH by foreign armed group	2.55** (1.01)	-0.97 (1.77)	0.79 (0.92)	-16.92 (10.83)
Past attack * Weak state		7.12*** (2.41)		31.64** (17.94)
Weak state		3.82*** (0.75)		53.54*** (9.99)
Past attack * Shabunda shock			10.40*** (3.75)	
Shabunda shock			12.01*** (1.52)	
F-stat				19.51
Control mean	1.89	1.58	1.23	1.58
Obs.	25,060	17,801	17,801	17,801

Notes: This table presents OLS estimates of equation 1 for attacks interacting with weak state indicators. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variable is an indicator for whether the respondent’s household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . Column (2) uses as weak state indicator if no state forces are present in the village in year t . Column (3) uses as weak state indicator whether respondent lived in Shabunda during 2003–05 and 2011–13 when military forces retreated from the area and created state weakness. Column (4) instruments the first weak state indicator (and the interaction term) using the second weak state indicator (and the interaction term). All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t and state force is present. P-value: *** 0.01, ** 0.05, * 0.10.

Table VIII: Summary Table: Parochial Altruism vs. Economic Motives

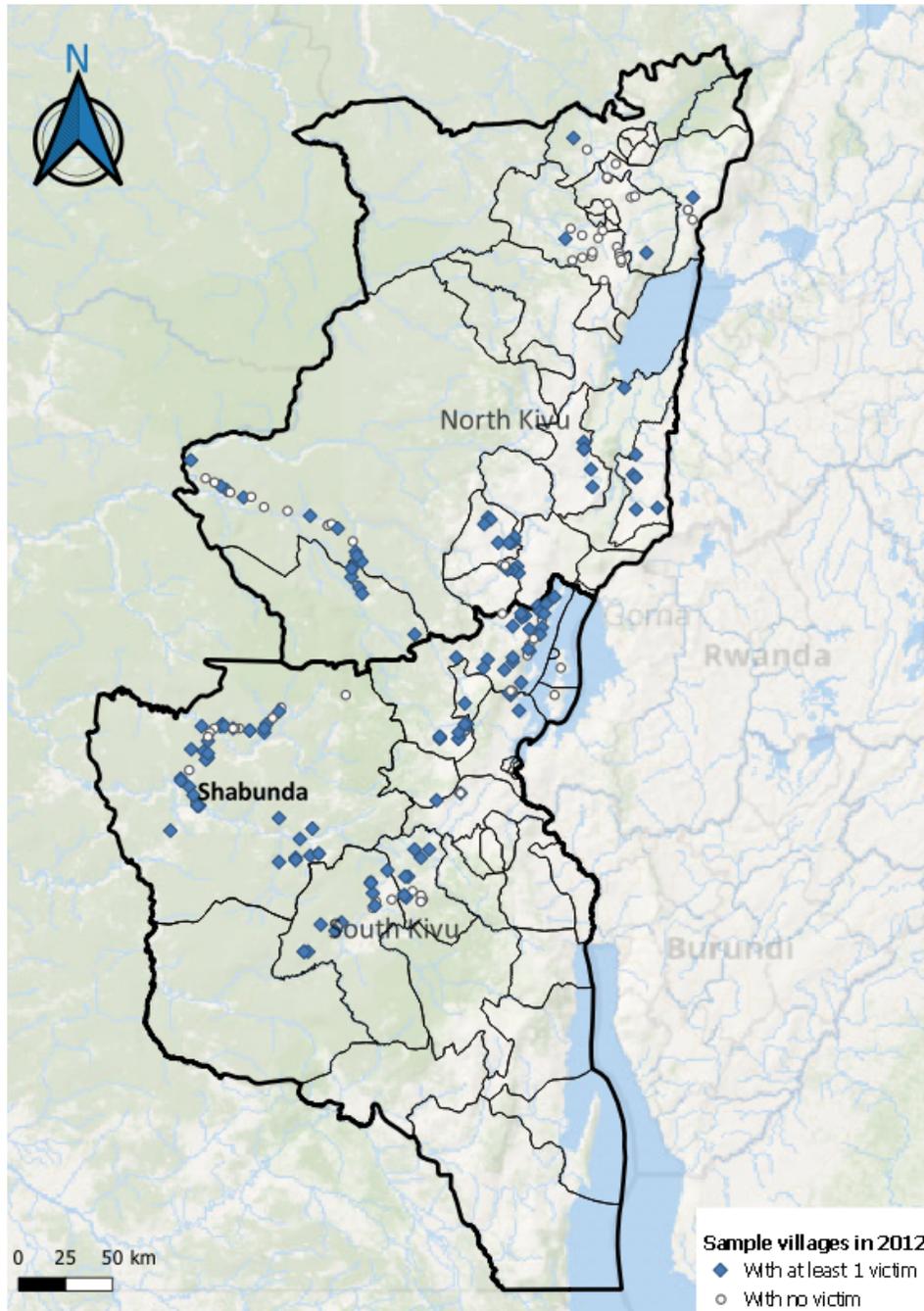
	Participation Militia (%)
<i>Control mean:</i>	
Never experienced an attack on HH by foreign armed group	1.89
<i>Effect of experienced shock:</i>	
An attack on HH by foreign armed group before year t	5.07** (2.01)
Local gold price increased in year t (Use world gold price as an IV)	
(a) by \$1 per g.	-0.21* (0.11)
(b) by 1 s.d. of local gold price	-1.70* (0.91)
(c) by 100% (using $\log P_t$)	-5.10* (2.72)
\Rightarrow One past attack on HH by foreign armed group has an equivalent effect of an decrease in local gold price by:	\$24.1 per g.
Daily production of gold miner (Geenen 2013)	1 g
Total number of work days per year (Assumed)	300 d
Tax by local authority (Own data)	45%
\Rightarrow Decrease in yearly income by:	\$3,977
GDP per capita in 2005 in DRC (World Bank)	\$218
\Rightarrow Decrease in yearly income as in GDP per capita:	18.2 times

Notes: This table compares the effect of past foreign armed group attack on household to the effect of gold price shock on participation in a Congolese militia. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year, residualized after controlling individual FE, village FE, year FE, and age FE. The explanatory variable are (i) an indicator for whether the respondent's household has been attacked by foreign armed groups and (ii) local gold price in a given year, instrumented by international gold price, all residualized after controlling the same set of fixed effects. The regression clusters two-way at the individual respondent and the village*year levels. We compute the effect of increase in local gold price by 1 s.d. (Row (b)) simply by scaling the coefficient in Row (a) by s.d. of local gold price (\$8.3 per gram). We compute the effect of increase in local gold price by 100% (Row (c)) by replacing gold price level with log of gold price in the regression. Coefficient of past foreign armed group attack on household remains almost the same in both specifications. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t and state force is present. P-value: *** 0.01, ** 0.05, * 0.10.

Figures

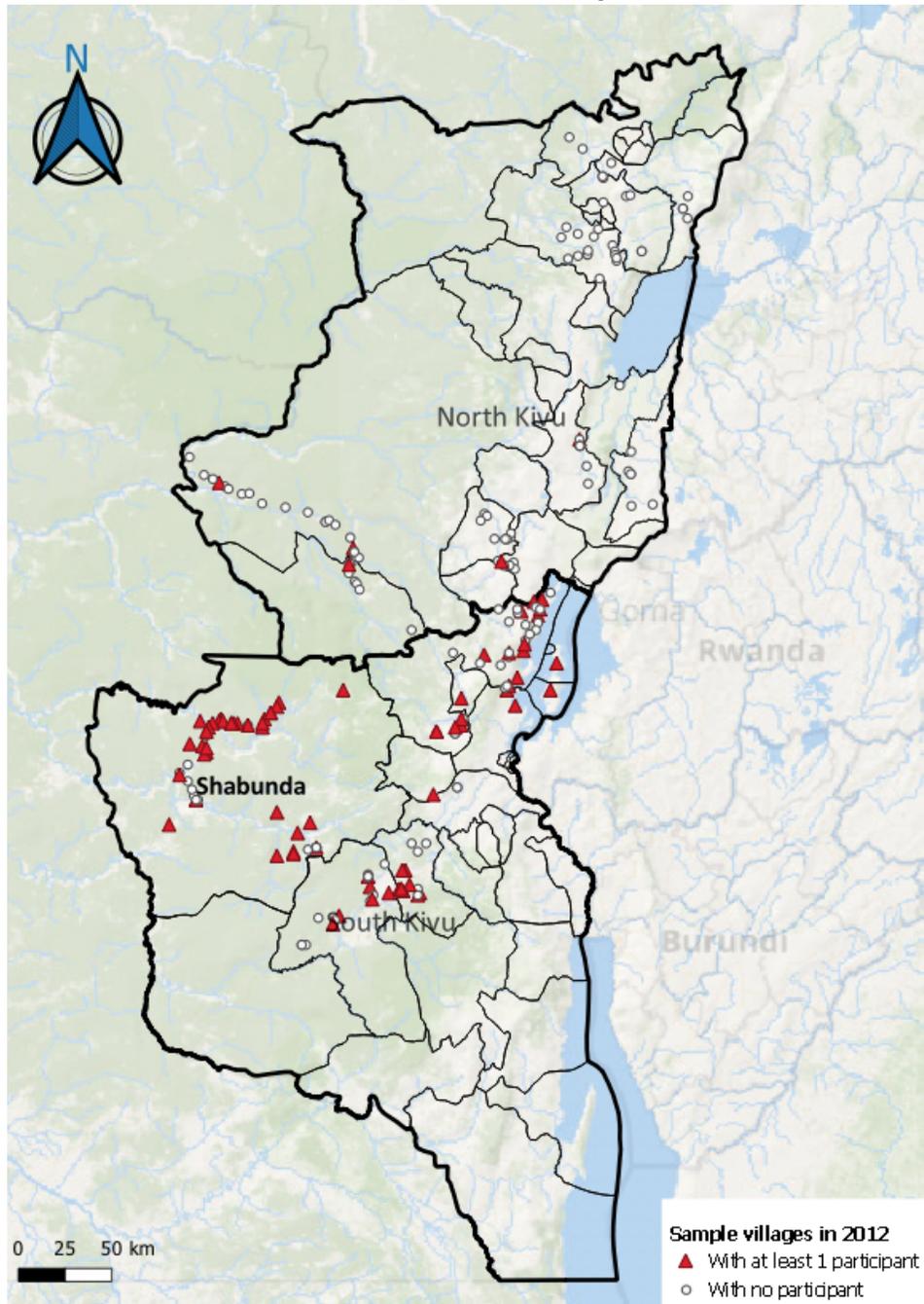
Figure I: Attacks and Participants in Militia, Spacial Distribution

Panel A: Attacks on sample households



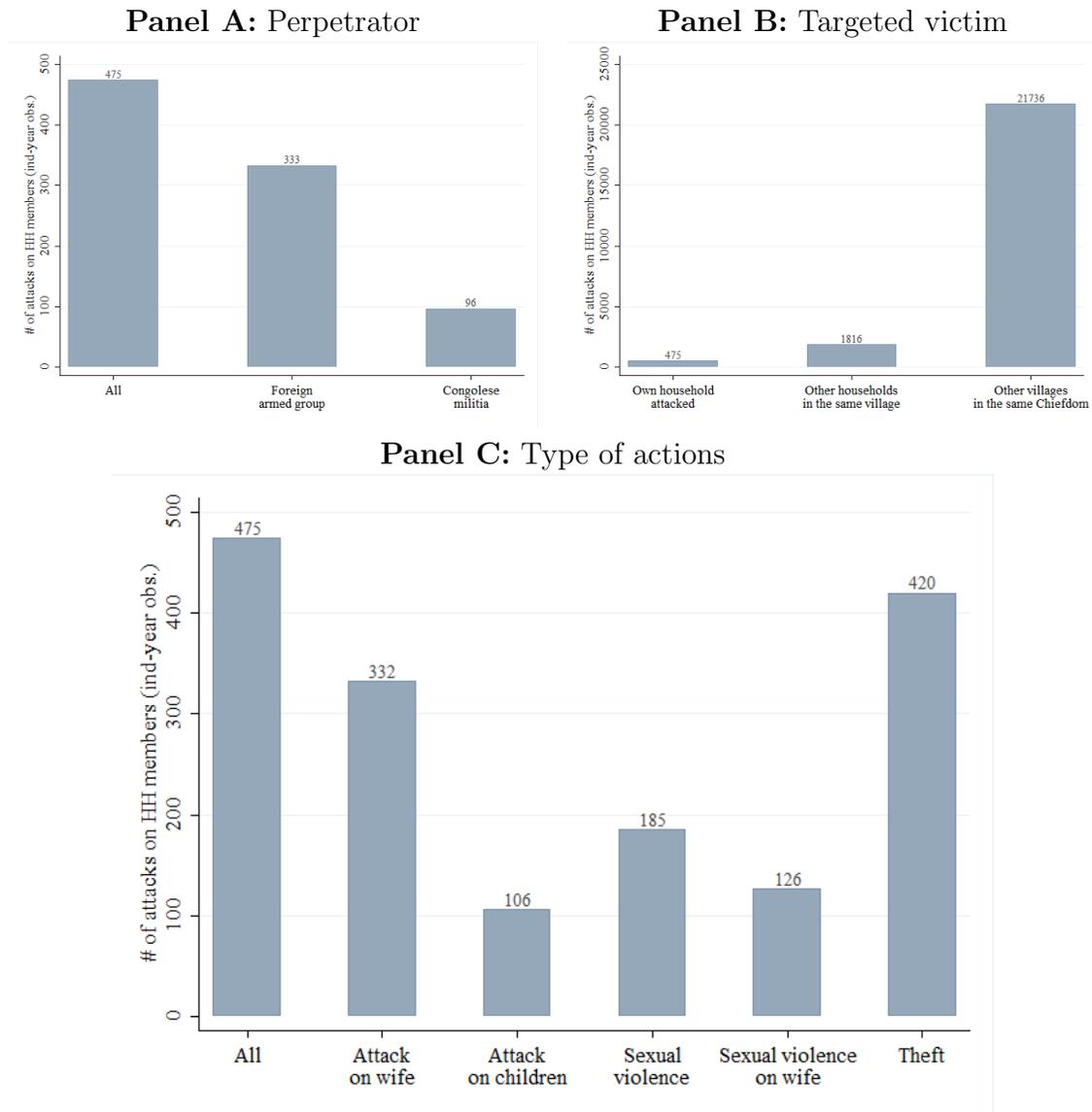
Notes: This figure shows spacial distribution of sample villages in year 2012. Blue diamonds are villages in 2012 where at least one respondent has experienced an attack on household up to 2012.

Panel B: Participants in a Congolese militia



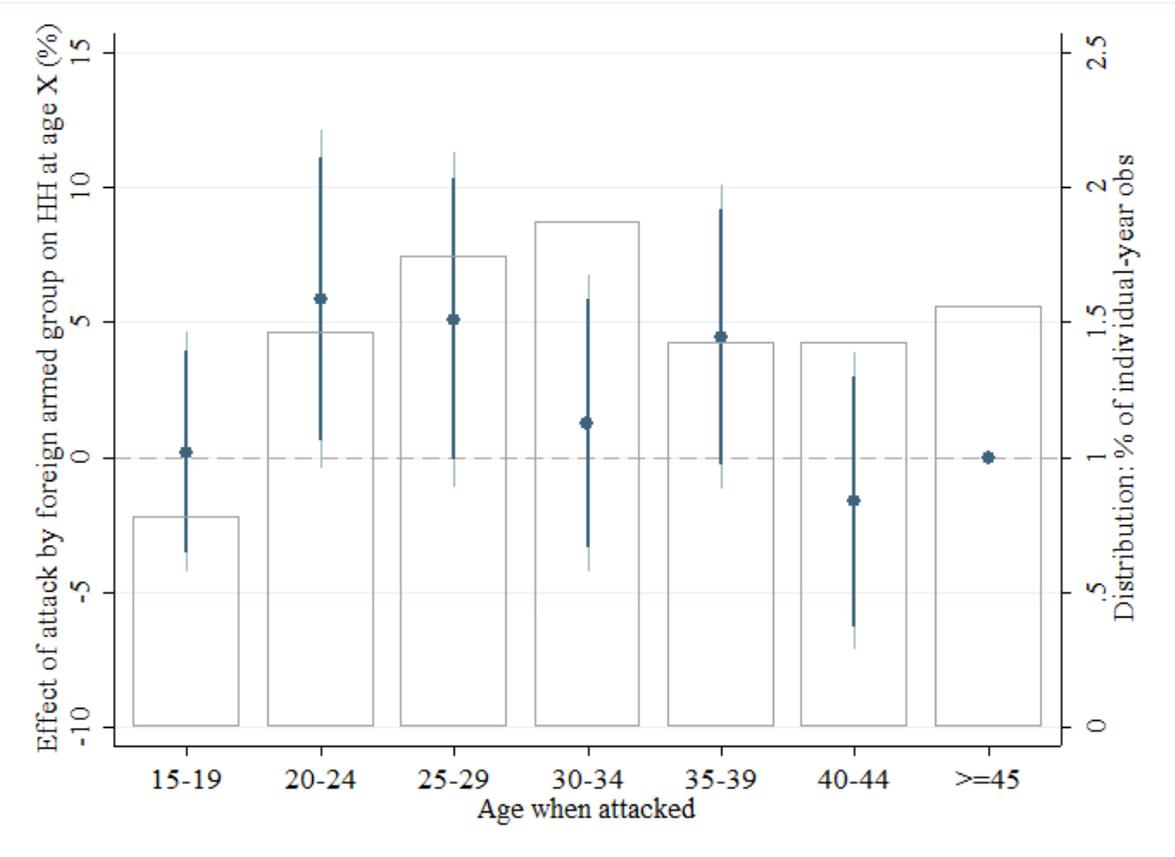
Notes: This figure shows spacial distribution of sample villages in year 2012. Red triangles are villages in 2012 where at least one respondent has participated in any armed group up to 2012.

Figure II: Attack Characteristics



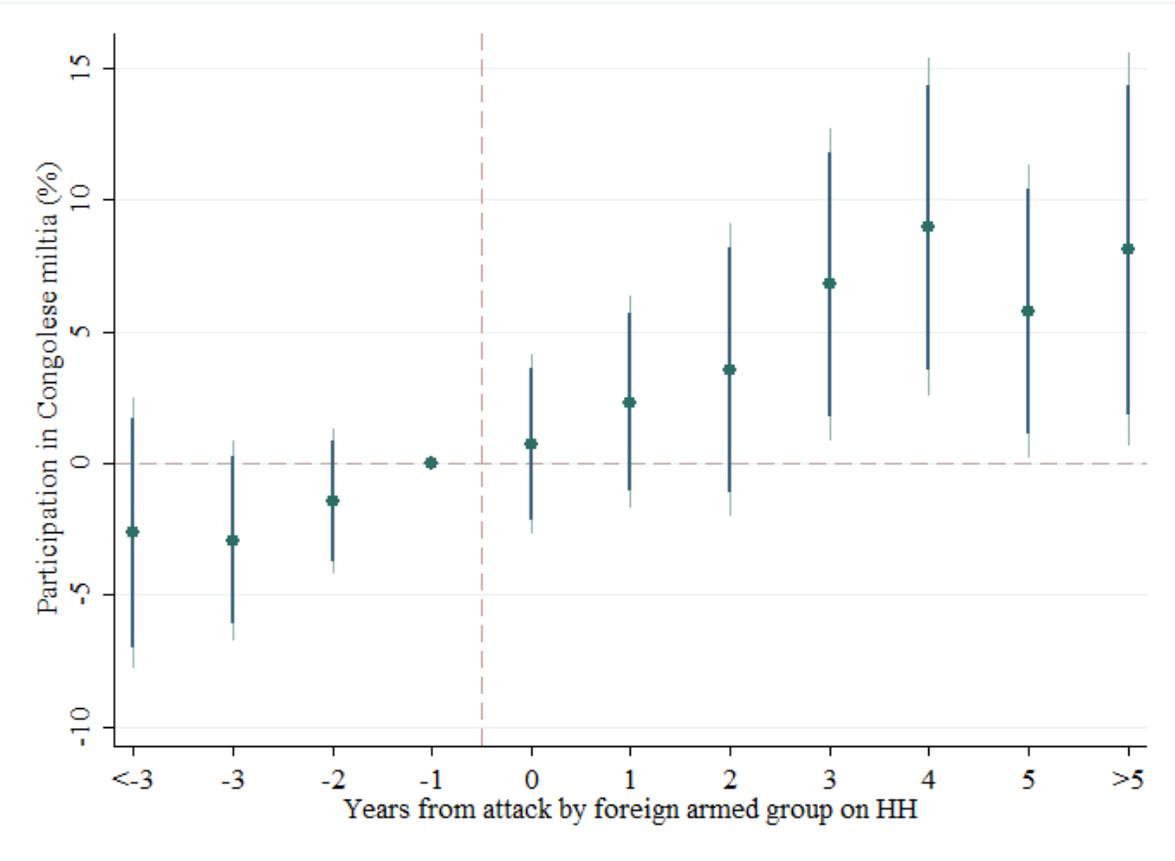
Notes: This figure shows the distribution of village attack observations with different characteristics. Panel A uses the sample of 475 reported attacks that targeted the households of the respondents in the sample and shows their distribution by perpetrator. Panel B uses the whole sample of attacks to have happened in each village reported by respondents and other households of the sample in the villages in the same Chiefdom. Based on this information, it shows the fraction of individual-year observations in which the own household was attacked, other households in the village were attacked, or other villages in the Chiefdom were attacked. Panel C decomposes all attacks on the household by the type of actions that were conducted (not mutually exclusive), respectively: attack on the spouse, attack on children, attack with sexual violence, attack involving sexual violence on respondent's spouse, attack in which household property was stolen. The figure does not show events of sexual violence directed against children, which amount to 30 cases.

Figure III: Formation of Preferences: Traumatic Events—Attacks Suffered at an Early Age



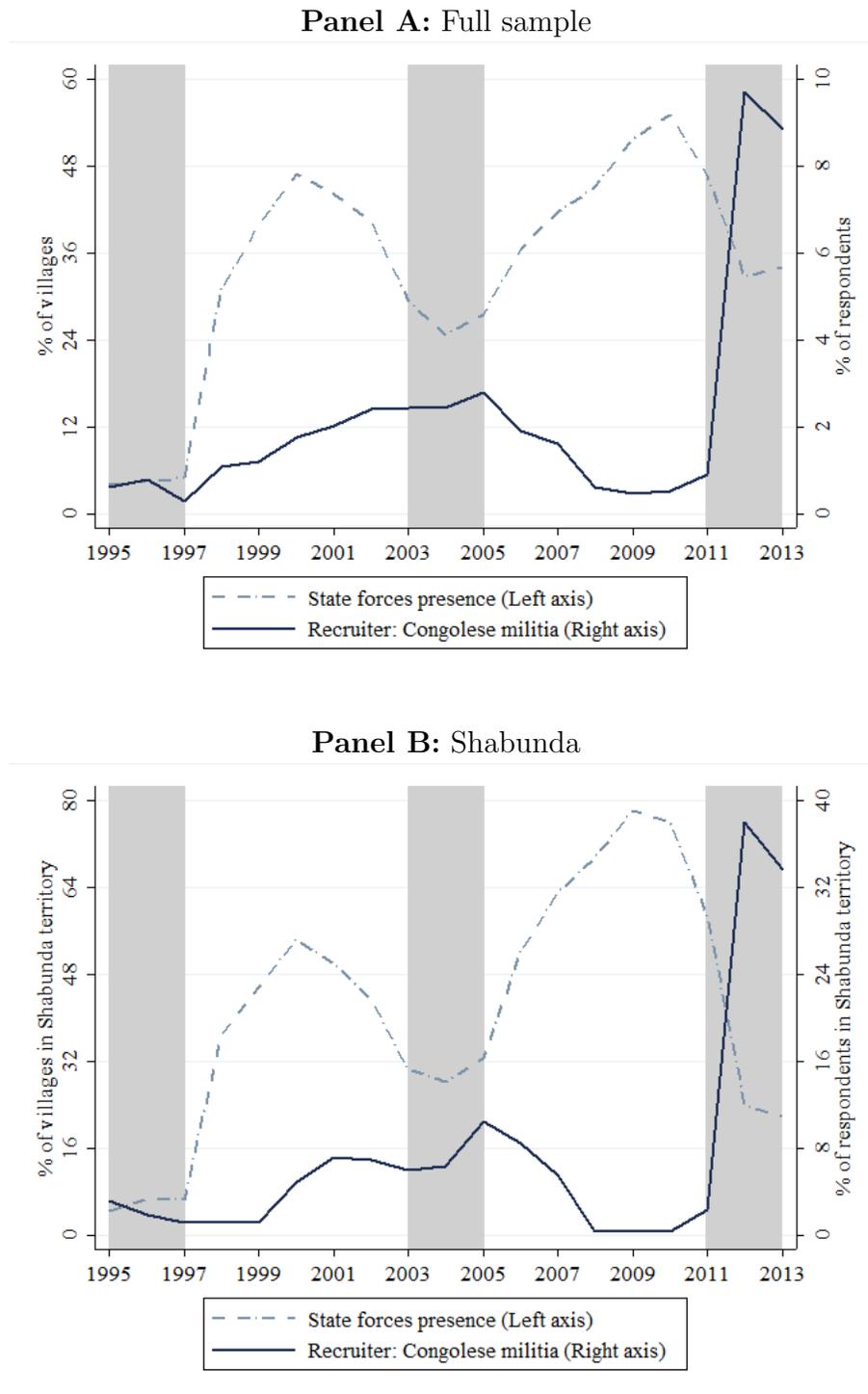
Notes: This figure estimates equation 1 by different age when attack happened on participation in a Congolese militia. All coefficients are estimated in one regression. The bars show the fraction (in percentage) of individual-year observations in which respondents have experienced a foreign armed group attack on household in each corresponding age group. We include observations between 1995 and 2013 above age 15 at year t and exclude individuals with no record of foreign armed group attack on household throughout the whole period. We do not include in the regression whether respondents experienced an attack before age 15 because this variable has no variation within individual when we only include observations above age 15. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. We show 95% and 90% confidence intervals for each coefficient.

Figure IV: Formation of Preferences: Effect Stability—Event Study Analysis



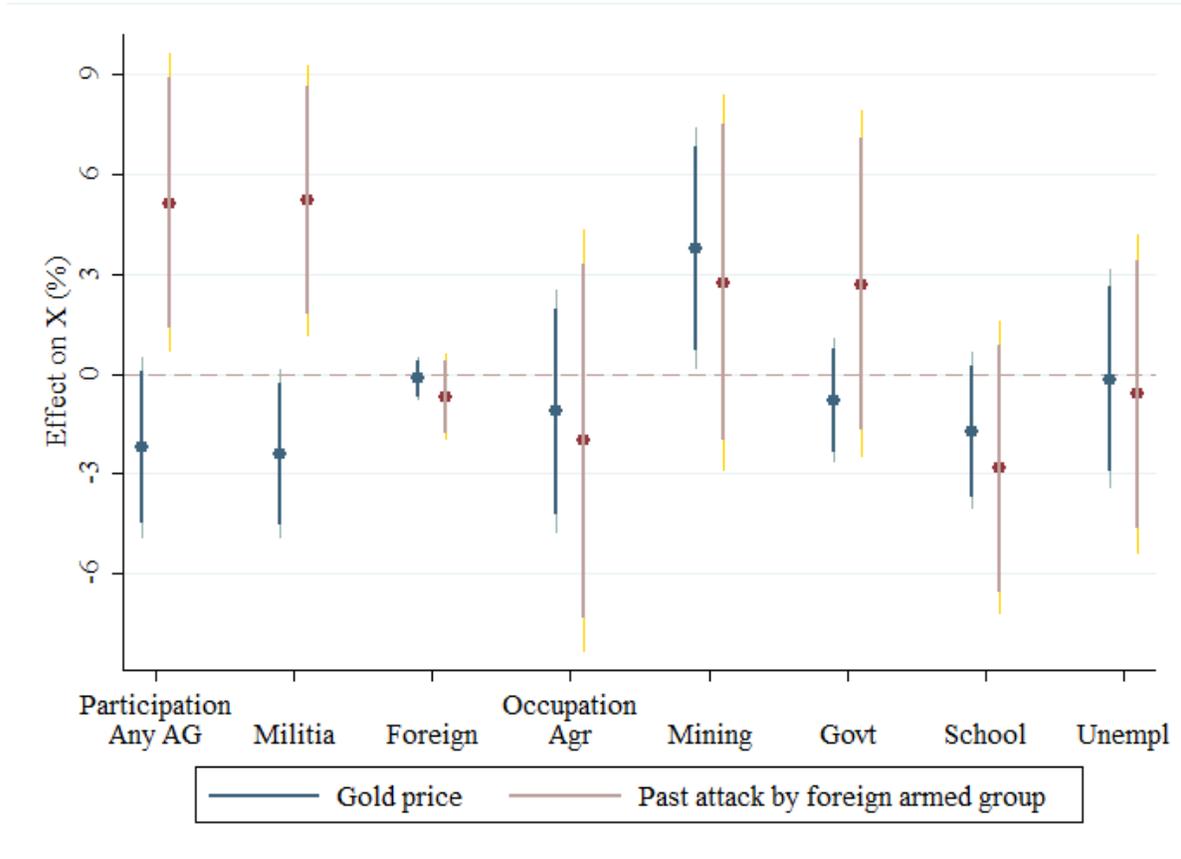
Notes: This figure shows the event study analysis of equation 1. We group lags after 5 years as an indicator of foreign armed group attack on household more than 5 years ago, and leads after 3 year as an indicator of foreign armed group attack on household more than 3 years after. We include observations between 1995 and 2013 above age 15 at year t and exclude individuals with no record of foreign armed group attack on household throughout the whole period. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. We show 95% and 90% confidence intervals for each coefficient.

Figure V: Formation of Preferences: Effect Stability—Interaction with Weak State Shocks



Notes: This figure shows the yearly trend of participation in a Congolese militia (blue solid line) and state force presence (green dashed line). State force includes Congolese national army and one particular foreign armed group (RCD) who took control of eastern Congo during Second Congo war. Grey areas refer to three state vacuum periods: pre-First Congo war (1995–97), post-Second Congo war Sun City agreement (2003–05), Kimia II military operation (2011–12).

Figure VI: Reduced Form Coefficient of the World Price of Gold on Occupational Choice



Notes: This figure shows the analysis of equation 2. Blue dots show the coefficient on world gold prices interacted with gold endowment, with different participation indicators and occupational choices as dependent variables. Red dots show the effect of past attack on household by foreign armed group on the same set of dependent variables. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. We show 95% and 90% confidence intervals for each coefficient. Table C.5, Panel A, presents this result in a table.

Online Appendix

A Data Description

A.1 Original questions on individual attacks

Variable	Survey question	Code
Module: Respondent attack history (Up to 9 attack events)		
A_{jt}	<i>Was there any violent event in village j where you lived in year t?</i>	= 1 if resp. reported a violent event in village j in year t
$Perp_{jt}$	<i>Who was the perpetrator?</i>	= F/M/CA if resp. reported a violent event where perpetrator was a foreign armed group, Congolese militia, or Congolese national army
Mot_{jt}	<i>What was the attack motive?</i>	= p/s/c if resp. reported a violent event where the motive was pillage, sanction, or conquest
A_{ijt}	<i>Were you physically assaulted during the attack?</i>	= 1 if resp. reported a violent event where resp. was physically assaulted
$Theft_{f(i)jt}$	<i>Was any property of your household stolen during the attack?</i>	= 1 if resp. reported a violent event where any property of his household was stolen
$Chief_{jt}$	<i>Was the village chief assaulted during the attack?</i>	= 1 if resp. reported a violent event where the village chief was attacked

Variable	Survey question	Code
Module: Household information		
$A_{f(i)t}$	<i>For each of your household members, including yourself, list three episodes he/she was assaulted</i>	= 1 if any of the household members (excluding resp. himself) reported being assaulted in year t
$Viol_{f(i)t}$	<i>For each of your household members, including yourself, list three episodes he/she was sexually victimized</i>	= 1 if any of the household members (excluding resp. himself) reported being sexually victimized in year t

Subscript j indicates that information comes from respondent attack module where respondents are asked about violent events in contemporary villages. The information can vary across different respondents who live in the same village in the same year, but for concise notation we do not add additional individual subscript. Subscript i indicates the action imposed on respondent i . Subscript $f(i)$ indicates the action was imposed on respondent i 's other household members, excluding respondent himself. In the next subsection, subscript o indicates the action was imposed on other households in the same village in year t .

A.2 Construction of main attack variables

Variable	Construction	Interpretation
$AA_{f(i)jt}$	$= A_{jt} \times (Mot_{jt} \neq c) \times A_{f(i)t}$	Whether resp. i reported an attack in year t in village j with nonconquest motive, and in which year any of the household members (excluding himself) reported being assaulted
I_{it}^{Victim}	$= \mathbb{1}(\exists t' < t, AA_{f(i)j(it')t'} = 1)$	Whether resp. i reported an attack on household before year t (j depends on i 's living history)

The main definition of attack in this paper focuses on reported violent events with nonconquest motives on other household members, excluding respondent himself. We consider violent events with conquest motives mainly involving combatants during war and less about civilians, and thus conquest motives do not capture the mechanism where civilians participate out of intrinsic preferences. We do not include attack on respondent himself because it might affect participation through additional mechanisms—that is, attack on respondent himself potentially affects the respondent’s capabilities, for instance, through handicap (see Section 6.2 and 8 for detailed discussion).

The main explanatory variable in specification 1, I_{it}^{Victim} , is constructed as an indicator whether respondent i reported any attack on household (excluding self) in the past. Subscript $j(it')$ stresses that reported attacks took place in villages where respondents lived in year t .

The fact that the main attack variable is constructed by combining information from different modules might complicate the interpretation in at least two scenarios in the following:

- Suppose a respondent reports two violent events in the same year, both with pillage motive. The first event was perpetrated by a foreign armed group, the second event was perpetrated by Congolese militia. The respondent also reports an attack on his spouse in the same year, and in reality his spouse was attacked in the second event. Our construction of main attack variable, however, would create a “false” attack on respondent’s spouse by a foreign armed group. This scenario, however, is unlikely to happen in our data. In total, 1,415 respondents have reported 4,097 nonconquest violent events, and 77.4% are reported in the year when the respondent does not report any other nonconquest violent events.
- Other household members might live in a different locations than the respondent in year t . This is also unlikely because the majority of the households observed in the data are nuclear family households. Out of 1,534 households that have detailed rosters of current family members that live with the respondent (the total number of households is 1,537), 68% of the households do not include family members other than spouse and children. If the respondent reported that his spouse or children were attacked in year t , we assume that his spouse or children were living with the respondent in the same village as well.

A.3 Asset variables

Each respondent in South Kivu is asked to list yearly purchase and sales for farm animals (cows, goats, and pigs) and fields since 1990. For asset stock at birth, we ask how many cows, goats, pigs, and fields the respondent’s father had when the respondent was born. We

also ask each respondent to report their asset stock at the survey year in farm animals but not fields.

We adopt the following approach to construct asset stock in cows, goats, and pigs. If respondent is not married, for farm animals and lands, we start from respondent's current asset stock and calculate respondent's asset stock in previous year by subtracting respondent's net purchase of asset this year from current asset stock. We calculate respondent's asset stock in each year backward up to year 1995. If respondent is married, we calculate respondent's asset stock backward up to the year when respondent was first married (89.9% of respondents who have hold marriages are only married once). Before the year respondent was first married, we start from respondent's asset stock at birth and calculate asset stock in following years by adding net purchase of asset up to the year before respondent was first married. The reason for this approach is that respondent is separated from his original household and starts a new household when he is married. For fields, since we did not ask each respondent to report their fields at the survey year, we calculate respondent's stock of fields starting from his stock of fields at birth and adding net purchase of fields in the years that follow. We further assume that when respondent is married, he gets one more piece of field in addition.

The construction of wealth variables above does not take into account the potential effect of attack on asset stocks. We use the following method to account for the loss of properties during a violent event. We first calculate the average loss in farm animals across all recorded violent events, and assume that each household would lose the average amount of farm animals if their household suffers from theft. Then, during the years when respondent reports a violent event with theft on the household, we decrease the total asset by the

assumed amount of loss of farm animals. We assume that violent events would not affect the stock of fields.

We then extract the principal component from the computed asset stock of cows, goats, pigs, and fields as the main wealth variable (see Table I, Table VI). The results are unchanged whether the calculation of the asset stocks account for loss of properties. For investment, we compute the principal component from the purchase of cows, goats, pigs, and fields. For the wealth of birth, we compute the principal component from the amount of cows, goats, pigs, and fields the respondent's father had at the respondent's birth, and the number of wives of his father and whether the respondent is a relative of the village chief.

B Additional background on FDLR and militia

B.1 Re-emergence of the Raia Mutomboki in 2011

On one side, a security vacuum was caused by the process of “regimentation” of the Congolese army (Stearns, 2013; Vogel, 2014). As in 2004–05, the FDLR took advantage of the security vacuum to take control in Shabunda. Under pressure from successive joint military operations by the Congolese and Rwandan military operations from 2009 to 2012 (in particular, the Umoja Wetu, Kimia II, and Amani Leo operations), the FDLR carried out violence against local populations, sparking the remobilization of the Raia Mutomboki. The 2011 movement, however, was considerably larger than the one in 2004–05, spreading to the northern areas of the territory of Shabunda in South Kivu, and eventually into the neighboring territories of Mwenga and Kalehe in South Kivu province, and parts of North Kivu province (Stearns, 2013; Hoffmann and Vlassenroot, 2014; Vogel, 2014).

B.2 Additional qualitative evidence on FDLR attacks

Consider the following account of a member of the Mera village:

Arriving [in the FDLR area], they accused us of being collaborators of the enemy.

We were immediately arrested and the decision was to kill us. We were beaten up, handcuffed, and they wanted to cut my head off. The machete was already on my neck and made an incision, when one of these Hutu had shouted to stop killing us because collaborating with the Tutsi was also their way of saving themselves from these people. We were released and returned to the village. When I arrived in

the village, I was dominated by anger and the only revenge and I decided to join the armed group with my brother [redacted] who was in [redacted] . . . I then stayed at least six years in this armed group and my exit or my demobilization was determined by the pacification of the country . . . It is through this process that I integrated civilian life and abandoned the army.

B.3 Additional Details on the Origins of the FDLR

The armed group known as the Front de Liberation du Rwanda (FDLR) is an ethnic Hutu group. In July 1994, a rebel movement took power in Rwanda, ending the genocide that had been perpetrated by government supported Hutu dominated militias, the Interahamwe, and the government forces, against the Tutsi. In response to the change of power, two million Rwandans, mostly Hutus, fled into eastern DRC, specifically North Kivu. Among them were the Interahamwe, but also former Rwandan state bureaucrats and armed forces. They formed the Armée de Libération du Rwanda (AliR), predecessor of the FDLR.

In 1996, the Rwandan government launched a military campaign that started the First Congo War (1996–97). One of the goals was to eliminate the insurgent threat coming from the Kivus. While the Rwandan coalition succeeded in defeating Congolese government forces, installing a new president, and occupying large parts of the country, they failed to completely defeat Rwandan rebel activity in eastern DRC.

Conflicts between the new Congolese government and its Rwandan and Ugandan backers in 1998 plunged the DRC into the Second Congo War (1998–2004). During this war, Rwanda backed a rebel group, the Rassemblement Congolais pour la Democratie (RCD), that quickly

controlled the eastern half of the country, where it overtook the apparatus of the state and all urban areas. In the countryside, resistance militia had formed, which the RCD fought through counterinsurgency campaigns. The Congolese state had no formal control over the east in this period (Verweijen and Vlassenroot, 2015, Clark, 2002, Ngonzola-Ntalaja, 2002).

Instead, the Congolese government supported various armed groups and provided them with funds and ammunition to fight the RCD. Among them were the former Rwandan government forces and militia members, AliR, who in 2000 formed the FDLR. The FDLR is, in most areas of DRC, a foreign-armed group. By 2004, all major armed groups, except the FDLR, vacated the east in exchange for benefits precluded in a peace agreement (Sun City peace agreement). The Congolese state struggled to regain control over the eastern provinces, creating a security vacuum. The FDLR took advantage, and expanded their territory in North and South Kivu.

Between 2004 and 2008, the FDLR became notorious as one of the most violent groups among a host of armed actors in the eastern DRC. The Rwandan government continued to support armed groups who fought against the FDLR, while the Congolese state alternatively tolerated or actively supplied the FDLR.

C Appendix Tables

Table C.1: Main Regression with Participation in Congolese Army and Past Attack by Congolese Army

Past attack on HH	Recruiter groups and identities					
	Congolese army		Congolese militia	Foreign armed group	Any armed group	
	(1)	(2)	(3)	(4)	(5)	(6)
By foreign armed group	-0.07 (0.07)					
By Congolese militia		-0.01 (0.15)				
By Congolese army			-0.14 (0.09)	6.16 (6.11)	0.40*** (0.15)	6.49 (6.10)
Control mean	0.19	0.20	0.18	2.25	0.31	3.11
Obs.	25,060	25,060	25,060	25,060	25,060	25,060

Notes: This table presents OLS estimates of equation 1, focused on Congolese army both as a recruiter and perpetrator. The dependent variable in columns (1)–(3) is an indicator for whether the respondent joins a Congolese army. The explanatory variable in each line is an indicator for whether the respondent’s household has been attacked by foreign armed groups, Congolese militia, and Congolese army, respectively. The dependent variables in columns (4)–(6) are indicators for whether the respondent joins a Congolese militia, foreign armed groups, or any armed group, respectively. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household by the corresponding armed group before year t . P-value: *** 0.01, ** 0.05, * 0.10.

Table C.2: Balance between Victims vs. Nonvictims

	HH never attacked (1)	HH attacked by foreign group		
		All (2)	Intention Pillage (3)	Intention Not pillage (4)
Obs. (All ages, 1995 $\leq t \leq$ 2013)	26,786	333	296	37
<i>Panel A. Socio-demographic background per individual-year obs</i>				
Age in year t	28.11	34.77***	35.01***	32.86***
Married in year t (%)	37.28	41.44**	42.23*	35.14
Works primarily in mining in year $t - 1$ (%)	11.31	14.44	14.84	12.50
Works primarily in agriculture in year $t - 1$ (%)	37.94	57.22***	56.13***	62.50***
Works primarily as civil servant in year $t - 1$ (%)	5.49	6.95	7.10	6.25
Works primarily in school in year $t - 1$ (%)	16.99	8.56***	9.68*	3.13***
Unemployed in year $t - 1$ (%)	28.27	12.83***	12.26***	15.63***
<i>Panel B. Economic status per individual-year obs</i>				
Wealth at birth (z-score, only $t = 2012$)	-0.04	0.18**	0.16**	0.25**
Asset stock in year $t - 1$ (z-score)	-0.03	0.05**	0.04**	0.09
Investment in year $t + 1$ (z-score)	0.02	0.42***	0.44***	0.28
<i>Panel C. Participation history per individual-year obs</i>				
Ever participated in any armed group before year t (%)	6.16	11.11	10.81	13.51
Ever participated in Congolese militia in year t (%)	3.98	6.91	6.76	8.11
Ever participated in foreign armed group in year t (%)	1.67	3.60	3.38	5.41

Notes: This table shows descriptive statistics for observations where respondents' household members are attacked by foreign armed groups versus not in year t . Column (1) shows the mean characteristics of observations in which the household is not attacked by foreign armed groups and column (2) shows those for individual-year observations in which the household is attacked by foreign armed groups. Stars indicate the levels of statistical significance for the difference between the means in column (1) and (2) including village and year fixed effects to isolate within village targeting, with standard errors clustered two-way at the individual respondent and the village*year levels. We examine balance on socio-demographic characteristics in panel A. Attacked individuals tend to be 7 years older at the time of the attack than nonattacked observations, 4 pp. more likely to be already married, 19 pp. more likely to work in agriculture and 15 pp. less likely to be unemployed. This suggests that attackers target more economically active and established households within village. To explore this possibility, Panel B shows balance on economic outcomes. We find that the wealth at birth indicator for attacked households in attacked years is 22% of sd. larger than nonattacked observations. Our imputed asset stock index is somewhat larger for attacked households, and investment in the following year is also significantly larger. Taken together, this suggests that richer households are more likely to be targeted. The last panel shows the means of past participation history of the respondents. We find that attacks do not target households with more history of participation in any armed group. This rules out that attackers target potential participants.

Table C.3: Examining Potential Confounders—ACLEd

	Attack radius									
	5 km (1)	10 km (2)	15 km (3)	20 km (4)	25 km (5)	30 km (6)	35 km (7)	40 km (8)	45 km (9)	50 km (10)
<i>Panel A. Dependent variable: past attack on HH by foreign armed group (%), OLS</i>										
ACLEd attack by foreign armed group	2.52 (1.75)	0.20 (1.12)	0.91 (0.99)	0.37 (0.86)	0.20 (0.84)	0.44 (0.78)	1.41* (0.73)	0.46 (0.72)	0.43 (0.73)	0.04 (0.73)
Control mean	10.42	10.22	9.71	9.46	9.24	8.60	7.28	7.40	7.22	6.93
Obs.	16,503	16,527	16,604	16,678	16,728	16,829	16,872	16,886	16,928	16,947
<i>Panel B. Dependent variable: participation in Congolese militia (%), OLS</i>										
ACLEd attack by foreign armed group	1.58 (1.58)	-1.04 (0.94)	-0.30 (0.88)	-1.14 (0.89)	-0.97 (0.86)	0.71 (0.82)	1.46* (0.86)	1.35 (0.84)	1.44* (0.86)	2.11*** (0.80)
Control mean	3.21	3.41	3.60	3.94	4.04	3.34	3.31	3.41	3.31	2.79
Obs.	16,503	16,527	16,604	16,678	16,728	16,829	16,872	16,886	16,928	16,947

Notes: This table presents additional robustness check of the effect of past attack on participation, using public data from the Armed Conflict Location Event Data Project (ACLEd). Panel A regresses the indicator for whether the respondent's household has been attacked by foreign armed groups on whether an attack by foreign armed group was reported (within a radius of X km around village j) in ACLEd before year t . Panel B regresses participation in a Congolese militia on whether an attack by foreign armed group was reported in ACLEd before year t . We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t . P-value: *** 0.01, ** 0.05, * 0.10.

Table C.4: Examining Alternative Causal Channels—Additional Table

	Investment (z-score $\times 100$) (1)	Overreporting any violent event (%) (2)
Past attack on HH by foreign armed group	-8.82 (7.17)	-3.73*** (1.17)
Control mean	6.82	3.89
Obs.	14,874	25,060

Notes: This table presents OLS estimates of the effect of past attack on alternative mechanisms. The explanatory variable is an indicator for whether the respondent’s household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . Dependent variable in Column (1) is current investment z-score. Dependent variable in Column (2) is whether respondent overreports any violent event on village in year t . Overreporting is defined as whether respondent reports a violent event on village in year t , but less than half of the other contemporary villagers observed in the sample report so and there is no corresponding record in village chief survey. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced a corresponding type of attack on household before year t . P-value: *** 0.01, ** 0.05, * 0.10.

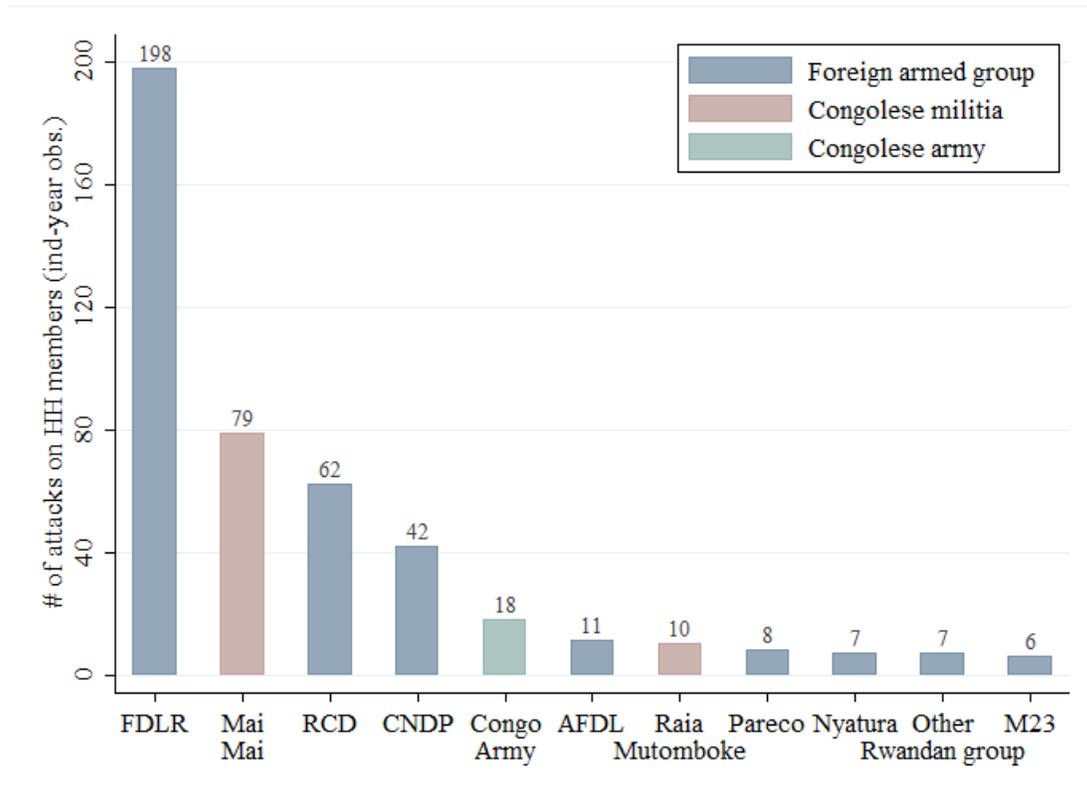
Table C.5: Economic Motives for Becoming a Rebel as Benchmark—Price Shocks

	Participation in t (%)			Occupation in t (%)				
	Any AG (1)	Militia (2)	Foreign (3)	Agr (4)	Mining (5)	Govt (6)	School (7)	Unempl (8)
<i>Panel A: Only main occupation and Sud Kivu observations</i>								
Gold $_j \times \log(P_t)$	-2.22 (1.39)	-2.41* (1.29)	-0.13 (0.33)	-1.13 (1.88)	3.79** (1.86)	-0.78 (0.96)	-1.73 (1.21)	-0.15 (1.68)
Past attack on HH by Foreign armed group	5.14** (2.29)	5.22** (2.08)	-0.70 (0.66)	-2.01 (3.24)	2.75 (2.90)	2.71 (2.66)	-2.84 (2.26)	-0.61 (2.44)
Control mean	4.92	4.24	0.30	53.64	11.10	6.59	8.27	20.40
Obs	12686	12686	12686	11580	11580	11580	11580	11580
<i>Panel B: Including non-main occupations, both South Kivu and North Kivu observations</i>								
Gold $_j \times \log(P_t)$	-0.92 (1.02)	-1.29 (0.93)	0.14 (0.28)	-2.46* (1.49)	-0.05 (1.71)	-0.82 (1.02)	-1.95 (1.22)	-0.15 (1.68)
Past attack on HH by foreign armed group	3.53** (1.53)	3.31** (1.39)	-0.17 (0.45)	-4.55** (2.26)	0.88 (2.25)	1.58 (2.81)	-3.57 (2.36)	-0.61 (2.44)
Control mean	3.81	2.91	0.43	65.98	20.37	7.33	8.65	20.40
Obs	17578	17578	17578	16472	16472	11580	11580	11580

Notes: This table presents OLS estimates of equation 2. The dependent variables are (a) indicators for whether the respondent joins any armed group, a Congolese militia, or a foreign armed group, respectively, in a given year, and (b) indicators for whether the respondent works in agriculture, mining sector, government office, is still a student or unemployed. The explanatory variables are world (log) price for gold interacting with gold endowment of the village, and an indicator for whether the respondent's household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . Panel A includes only main occupations and only limits to observations in South Kivu. Panel B includes non-main occupations and North Kivu observations that have available data on some occupations or gold endowment. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced any attack on household by foreign armed group before year t and did not live in a village endowed with gold. P-value: *** 0.01, ** 0.05, * 0.10.

D Appendix Figures

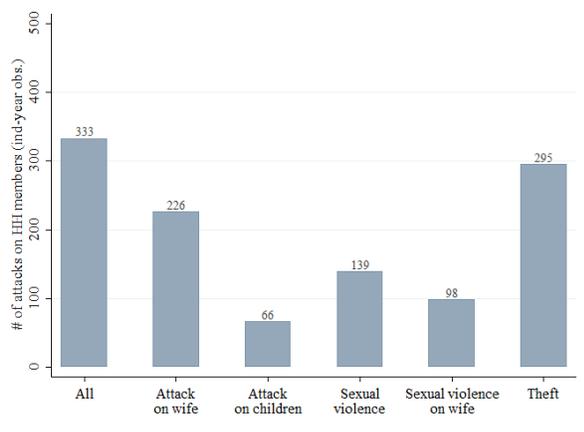
Figure D.1: Descriptive Summaries of Attacks by Perpetrator Identity



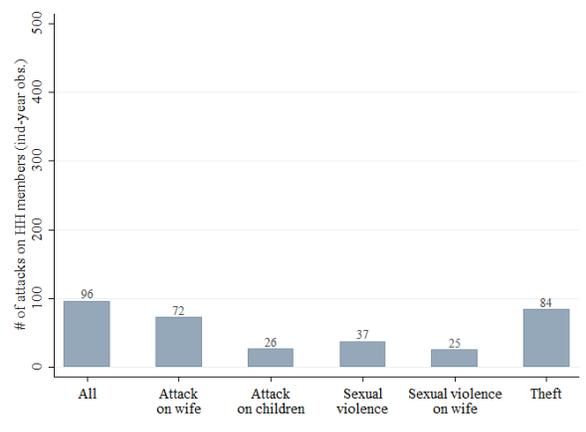
Notes: This figure uses the sample of 475 reported attacks that targeted a household of the respondents in the sample and shows their distribution by detailed perpetrators. Blue bars refer to foreign armed groups; Red bars refer to Congolese militia; Green bar refers to Congolese national army.

Figure D.2: Descriptive Summaries of Attacks by Type

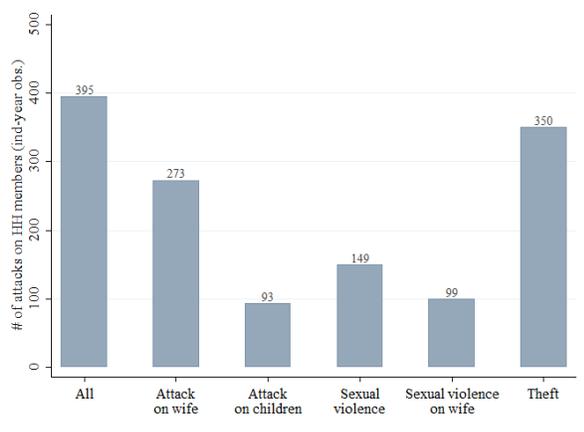
Panel A: Perpetrator: Foreign armed group



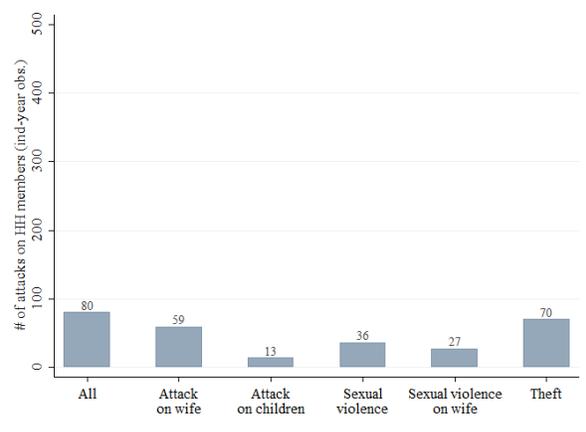
Panel B: Perpetrator: Congolese militia



Panel C: Intention: Pillage

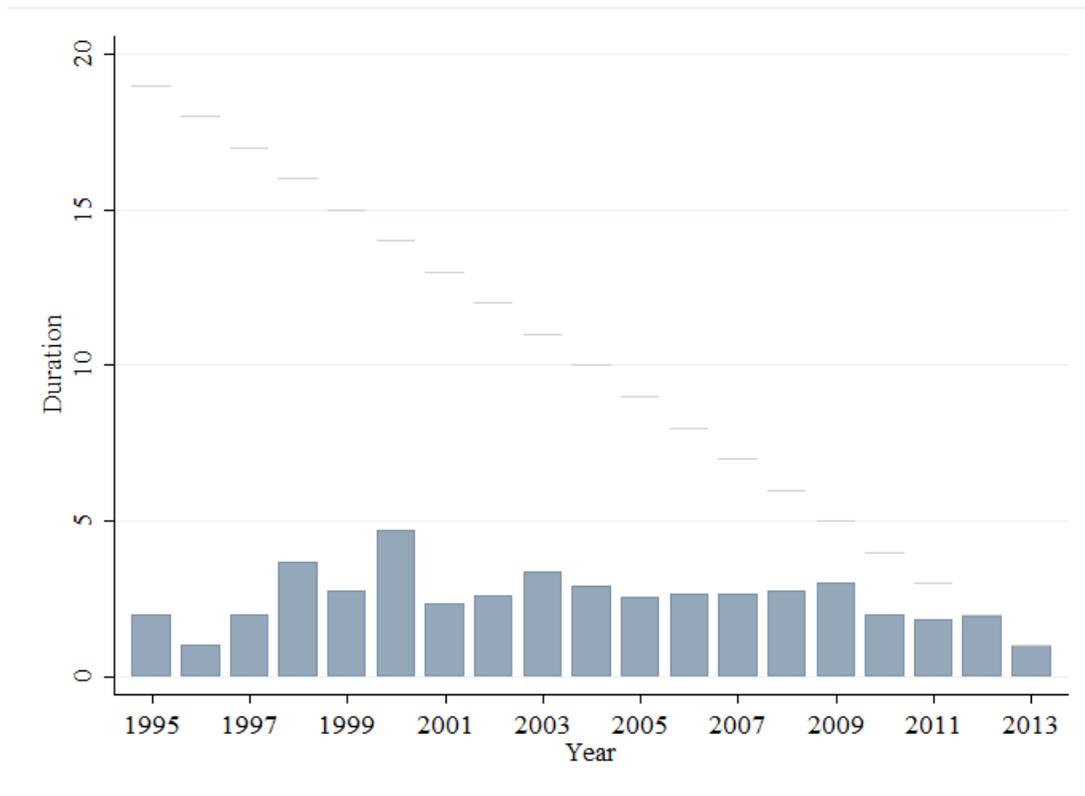


Panel D: Intention: Punishment



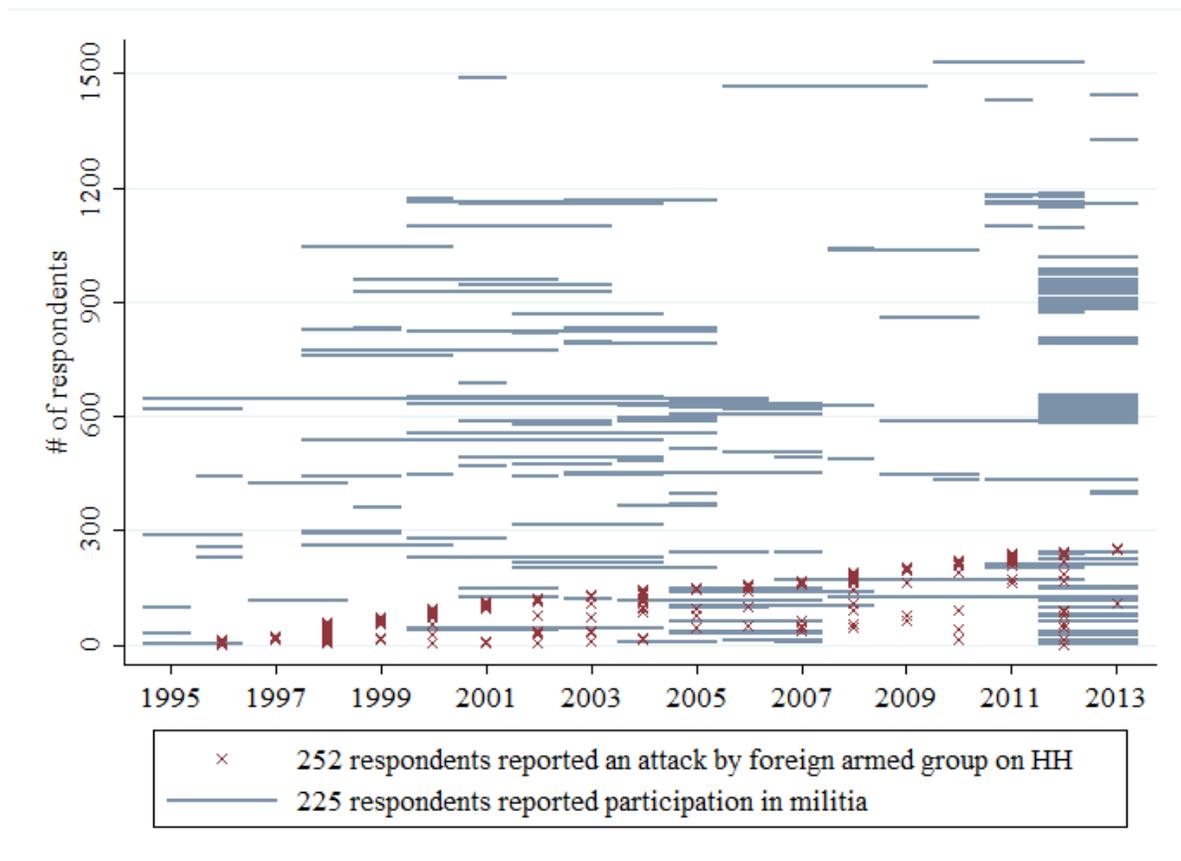
Notes: This figure decomposes different types of attacks on the household by the type of actions that were conducted (not mutually exclusive), respectively: attacks on the spouse, attack on children, attack with sexual violence, attack with sexual violence on spouse, attack in which household property was stolen. We look at four types of attacks respectively: attacks by foreign armed group, attacks by Congolese militia, attacks with intention to pillage, attacks with intention to punish civilians.

Figure D.3: Average Participation Duration in a Congolese militia



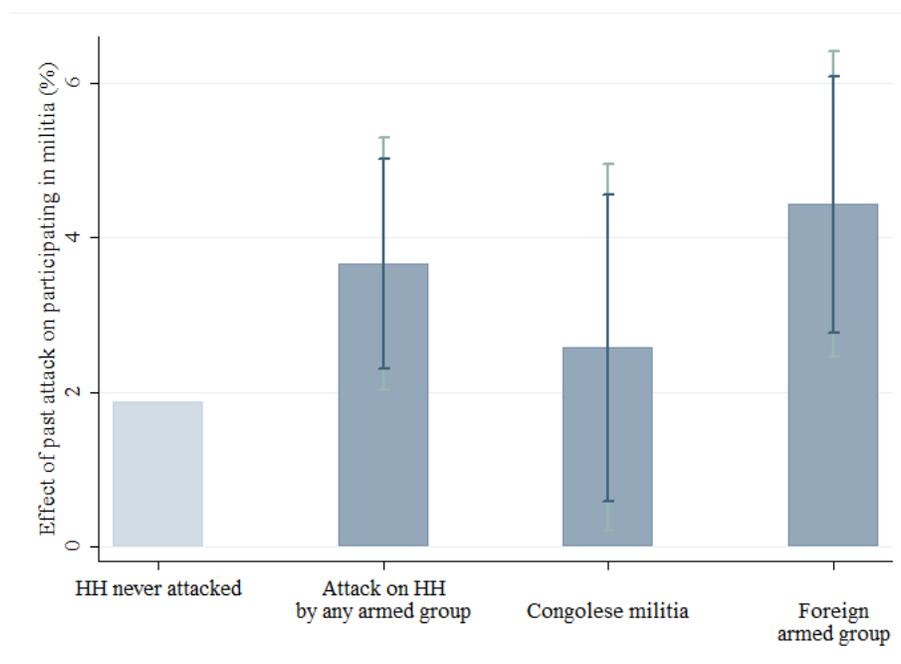
Notes: This figure shows average participation duration in a Congolese militia for each year since 1995. Gray line indicates the maximum year a participant can stay in an armed group.

Figure D.4: Attacks on Household Foreign Armed Group and Participation in a Congolese militia



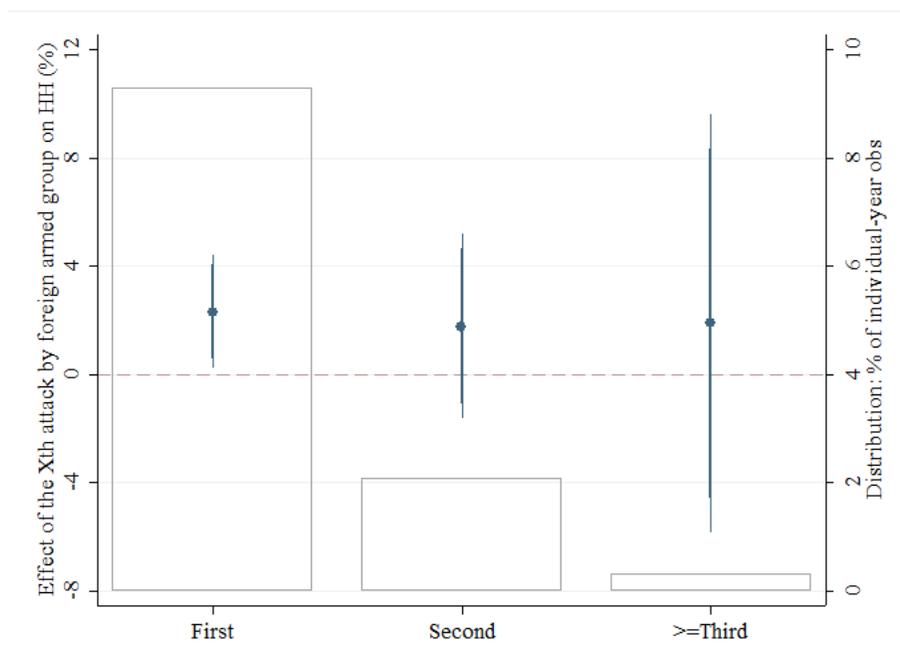
Notes: This figure presents all participation episodes in Congolese militia in the sample from 1995 to 2013 and foreign armed group attacks on the household. Y axis ranks individuals by their date of first reported attack by foreign armed group. X axis shows the years. Red crosses indicate the timing of an attack by a foreign armed group against the household for a given respondent in a given year. Blue lines indicate individual episodes of participation.

Figure D.5: Graphical Representation of the Main Result



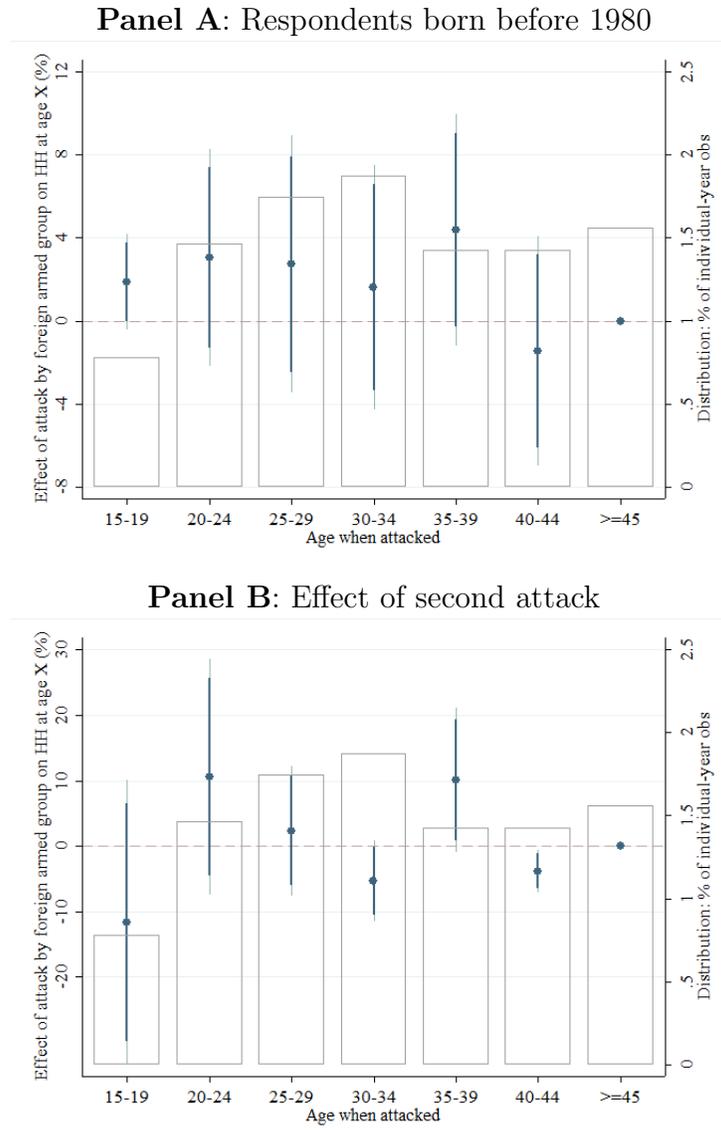
Notes: This table presents OLS estimates of equation 1 for different perpetrators (by any armed group, by Congolese militia, and by foreign armed groups respectively) and on participation in a Congolese militia. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variable in each bar is an indicator for whether the respondent's household has been attacked by any armed group, a Congolese militia, foreign armed groups, respectively. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household from any armed group before year t . 95% and 90% confidence intervals are shown. See Table II for more discussions.

Figure D.6: Order Analysis: The Main Result is Concentrated on the First Attack



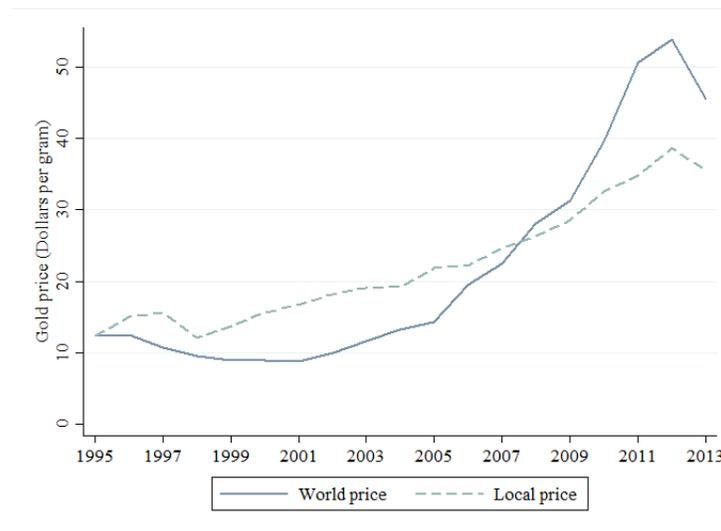
Notes: This table presents OLS estimates of equation 1 for attacks by foreign armed groups and on participation in a Congolese militia. The dependent variable is an indicator for whether the respondent joins a Congolese militia in a given year. The explanatory variables are whether the respondent's household has been attacked by foreign armed groups for the first, second, and third (and more) time. The bars show the fraction (in percentage) of individual-year observations in which respondents have experienced the first, second, and third (and more) foreign armed group attack on household. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. We show 95% and 90% confidence interval. See Table II for more discussions.

Figure D.7: Formation of Preferences: Role of Traumatic Events—Attacks Suffered at an Early Age, Additional Analysis



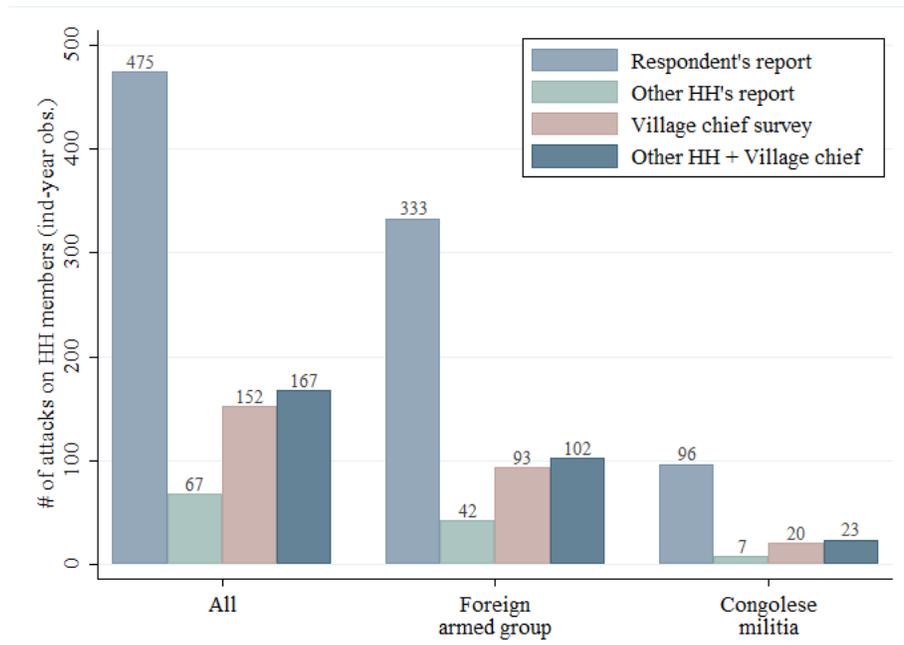
Notes: This figure replicates Figure III, with two deviations. Panel A shows the effect only for respondents who were born before 1980, and thus who are at least 18 years old during First and Second Congo wars and for whom participation into militia was an option. Panel B replicates Figure III using only the second attack. In each panel all coefficients are estimated in one regression. The bars show the fraction (in percentage) of individual-year observations in which respondents have experienced a foreign armed group attack on household at each corresponding age group. We include observations between 1995 and 2013 above age 15 at year t and exclude individuals with no record of foreign armed group attack on household throughout the whole period. We do not include in the regression whether respondents experienced an attack before age 15 because this variable has no variation within individual when we only include observations above age 15. All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Brackets show 95% and 90% confidence intervals for each coefficient.

Figure D.8: Times-series of World and Local Gold Prices



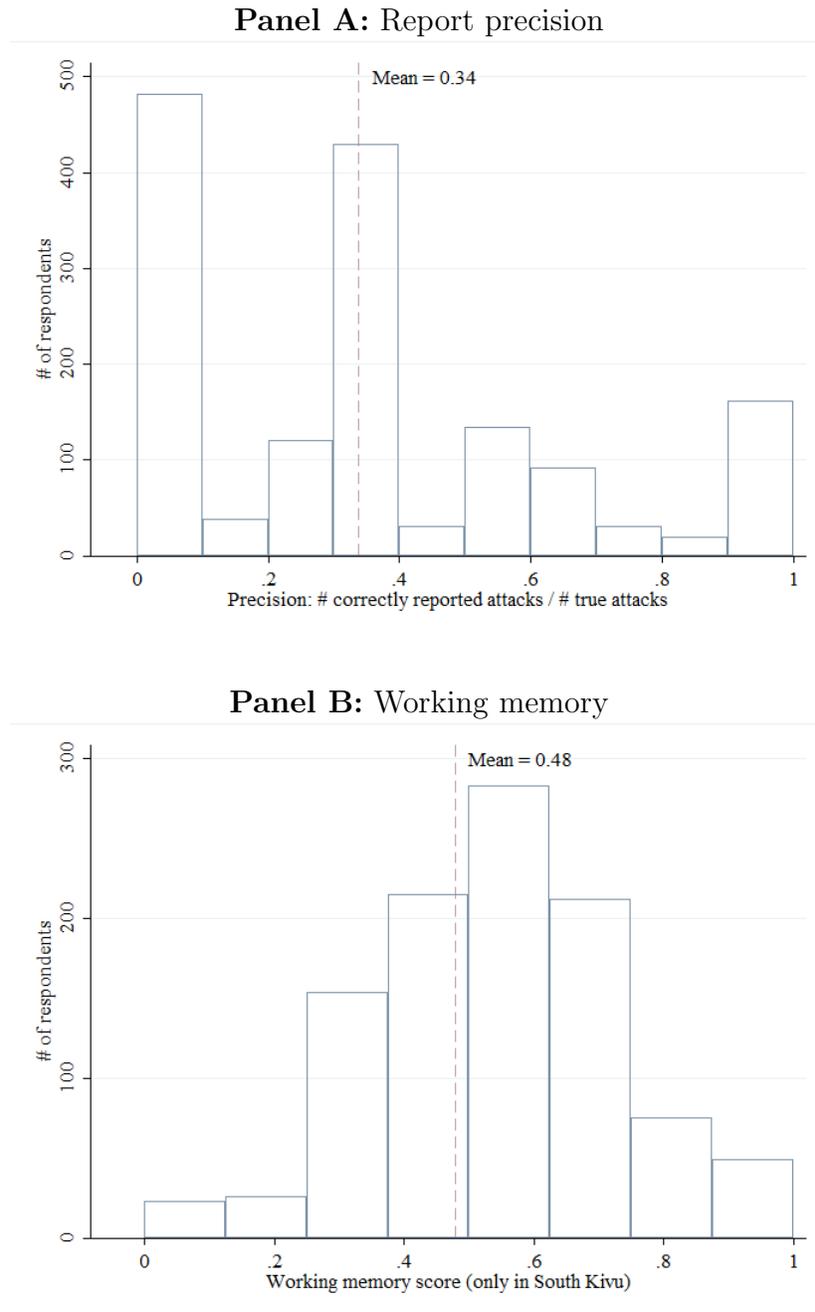
Notes: This figure shows the yearly world and local gold prices between 1995 and 2013.

Figure D.9: Cross-validation of Attack Variables



Notes: This figure shows the cross-validation of main attack variables. Light blue bars are reported attacks on household between year 1995 and 2013 from respondents' report. Green bars show attacks on household that are also reported by more than half of other respondents observed in the sample who lived in the same village in the same year. Red bars show attacks on household that are also reported in village chief survey. Dark blue bars show attacks on household that are cross-validated by both village chief survey and more than half of other contemporary respondents.

Figure D.10: Distribution of Report Precision and Working Memory



Notes: This figure shows the distribution of report precision and working memory which are used as weights in Table III, Panel B. Report precision in Panel A is defined as the proportion of all the “true” attacks, defined as attack episodes reported by each respondent and also by more than half of other contemporary respondents in the same village or by the village chief. The mean of report precision is 0.34. Working memory in Panel B is constructed from two memory tests described in Section 6.2. The mean of working memory test is 0.48.

E Mathematical appendix

E.1 Model

A unitary household makes decision for infinite time horizon. At time t , the household head decides (1) consumption and investment for all periods (but cannot commit to the plan), and (2) participation into an existing armed group for all periods.⁵¹ We denote C_{it} the consumption at time t for household i , and P_{it} the indicator of whether household head i decides to join an existing armed group at time t . The flow utility is defined as follows:

$$U(C_{it}, P_{it}) = u(C_{it}) + [\theta_{it}(A_{i0}, A_{i1}, \dots, A_{i,t-1}) - \phi_t(\bar{\theta}_t, M_t) - \zeta_{it}] \cdot 1(P_{it} = 1). \quad (3)$$

$u(C_{it})$ is a generic concave function satisfying Inada conditions. θ_{it} is the intrinsic utility of joining armed group, depending on household head's past attack history $A_{i0}, A_{i1}, \dots, A_{i,t-1}$, which can be interpreted as "revenge utility."⁵² ϕ_t is the cost of participation. It depends on two factors: (1) Exogenous constraint M_t . For example, if state force is present, villagers are discouraged to participate in militia. In the meantime, state force might recruit more villagers, so the effect of M_t on ϕ_t can be ambiguous. (2) Average revenge utility $\bar{\theta}_t$. We assume that if villagers enjoy higher utility from revenge on average, the psychological cost for participation is lower, which is designed to capture spillover effect. ζ_{it} is an idiosyncratic utility shock to participation, following some distribution $G(\cdot)$. Revenge (dis)utility can be realized only if the household head participates.

Attack: Attack is realized after household head makes participation decision. We assume

⁵¹We abstract from the framework endogenous formation of armed groups and leave it for future research.

⁵²We assume attack is a continuous variable for simple math derivation. Higher value of A_{it} can be also interpreted as more violence involved. We remain agnostic as to the effect of past attack on intrinsic utility.

each attack event is individually and independently distributed across all t and all i and follows some distribution $F(\cdot)$ with mean a_{it}^1 if household head participates or a_{it}^0 if not.⁵³

Death: We consider a possibility of terminating all flow utility after attack is realized. We assume the *expected* survival rate μ is a decreasing function of the attack likelihood a_{it} . For simplicity, we assume $\mu(a_{it}^1) - \mu(a_{it}^0) = \eta(a_{it}^1 - a_{it}^0)$ for some non-increasing function $\eta(\cdot)$.

Income: If household head survives, he earns labor income w through production, or T through appropriation by participating in armed group. We assume $T > w$.

Investment: Household invests the rest of the labor income on capital. Household possesses initial capital K_{it} . Capital depreciates at the rate of δ_{it} . Depreciation rate is higher if household suffers an attack: $\frac{\partial \delta_{it}}{\partial A_{it}} > 0$. Household earns capital gain r from each unit of current remaining capital.

Timing. (1) Household decides whether or not to participate, upon which intrinsic revenge utility θ_{it} , participation cost ϕ_t and idiosyncratic utility shock ζ_{it} are immediately realized. (2) Attack A_{it} is drawn from $F(\cdot | a_{it}^0)$ or $F(\cdot | a_{it}^1)$ depending on participation status. (3) If household survives, then household decides consumption C_{it} and investment I_{it} . (4) Household enters the next period.

Proposition 1 characterizes the effect of past attack on consumption, capital level, and the present value of participation compared to nonparticipation.

Proposition 1. *Suppose attack A_{it} is a continuous variable, capital depreciation rate δ_{it} is differentiable with regard to A_{ik} , $\forall k \leq t$. Then consumption C_{it} and capital $K_{i,t+1}$ decrease*

⁵³We assume idiosyncratic attack likelihood for two reasons. First, individuals with different characteristics might be targeted differently; for instance, see Table C.2. Second, a_{it} can be interpreted as subjective expectation of attack likelihood, which varies for different individuals. We also remain agnostic whether $a_{it}^1 > a_{it}^0$ or not. If $a_{it}^1 > a_{it}^0$, then participation may lead to more targeting from enemies. If $a_{it}^1 < a_{it}^0$, then participation may provide more protection for the household.

in A_{ik} , $\forall k \leq t$, whether participating or not.

Proposition 2 decomposes the effect of past attack on participation into intrinsic motivation, wealth effect, and protection effect.

Proposition 2. *Suppose attack A_{it} is a continuous variable, capital depreciation rate δ_{it} , intrinsic utility θ_{it} , attack likelihood a_{it}^0 and a_{it}^1 , and survival function $\mu(\cdot)$ are differentiable with regard to A_{ik} , $\forall k \leq t$. Then the effect of A_{ik} on participation likelihood p_{it} can be decomposed into a linear combination of the following three components:*

1. *Intrinsic motivation: $\frac{\partial(\theta_{it}-\phi_t)}{\partial A_{ik}}$*
2. *Wealth effect, positive if marginal utility of contingent consumption weighted by attack likelihood is lower when participating;*
3. *Protection effect*

From the proof of Proposition 2, we are able to rule out wealth effect by controlling in the main regression investment proxy (in the empirical analysis, the principal component of investment in farm animals). Similarly, we can rule out protection effect by controlling subjective attack likelihood (in our empirical analysis, whether respondent overreports any violent event).

E.2 Solution

We solve this problem by writing Bellman equation first. Some notations for simplicity:

$V_t^{N,s}(K)$: Present value if not participating in period t , after household head survives

$V_t^{P,s}(K)$: Present value if participating in period t , after household head survives

$V_t^N(K)$: Present value if not participating in period t , before attack is realized

$V_t^P(K)$: Present value if participating in period t , before attack is realized

$V_t(K)$: Present value before household head makes participation decision

Let β be the discount rate, and suppose household survives. The continuation utility from not participating is as follows:

$$\begin{aligned} V_t^{N,s}(K_t) &= \max_{C_t, I_t} u(C_t) + \beta E_t[V_{t+1}(K_{t+1})|P_t = 0] \\ \text{s.t. } C_t + I_t &= w + r(1 - \delta_t)K_t, \quad K_{t+1} = (1 - \delta_t)K_t + I_t. \end{aligned}$$

The first order condition is:

$$u'(C_t^N) = \beta E_t[V'_{t+1}(K_{t+1}^N)], \quad (4)$$

where $C_t^N, K_{t+1}^N = (1 - \delta_t)(1 + r)K_t + w - C_t^N$ is the optimal choice given nonparticipation.

For the continuation utility from participation, similarly,

$$\begin{aligned} V_t^{P,s}(K_t) &= \max_{\{C_t\}} u(C_t) + \beta E_t[V_{t+1}(K_{t+1})|P_t = 1] \\ \text{s.t. } C_t + I_t &= T + r(1 - \delta_t)K_t, \quad K_{t+1} = (1 - \delta_t)K_t + I_t. \end{aligned}$$

The first order condition becomes:

$$u'(C_t^P) = \beta E_t[V'_{t+1}(K_{t+1}^P)], \quad (5)$$

where C_t^P , $K_{t+1}^P = (1 - \delta_t)(1 + r)K_t + T - C_t^P$ is the optimal choice given participation.

Let's move backward before attack is realized. Household head would weight the present value by survival likelihood:

$$\begin{aligned} V_t^N(K_t) &= \mu(a_t^N)V_t^{N,s}(K_t) \\ V_t^P(K_t) &= \theta_t - \zeta_t + \mu(a_t^P)V_t^{P,s}(K_t). \end{aligned}$$

And household chooses to participate when $V_t^P(K_t)$ is larger:

$$V_t(K_t) = \max[V_t^N(K_t), V_t^P(K_t)].$$

Hence, at the beginning of each period, the probability of household participating would be:

$$\begin{aligned} p_t(x) &\equiv Pr[V_t^P(x) > V_t^N(x)] \\ &= Pr[\zeta_t < \theta_t + \mu(a_t^P)V_t^{P,s}(x) - \mu(a_t^N)V_t^{N,s}(x)] \\ &= G\left[\underbrace{\theta_t}_{\text{Intrinsic utility}} + \underbrace{\mu(a_t^P)V_t^{P,s}(x) - \mu(a_t^N)V_t^{N,s}(x)}_{\text{Protection and wealth}} \right] \end{aligned} \tag{6}$$

E.3 Proofs

Before we prove the two propositions, we claim that there is no clear prediction on the effect of participation on consumption or investment. Since value function $V_{t+1}(\cdot)$ is the maximum of $V_{t+1}^P(\cdot)$ and $V_{t+1}^N(\cdot)$, the expectation of $V'_{t+1}(\cdot)$ depends on the likelihood of participating

in $t + 1$. Break down $E_t[V'_{t+1}(x)]$:⁵⁴

$$\begin{aligned}
E_t[V'_{t+1}(x)] &= E_t p_{t+1}(x) \frac{\partial E_t(V_{t+1}^P)'}{\partial x} + (1 - E_t p_{t+1}(x)) \frac{\partial E_t(V_{t+1}^N)'}{\partial x} \\
&= E_t p_{t+1}(x) E_t [\mu(a_{t+1}^P)(1 - \delta_{t+1})(1 + r)u'(C_{t+1}^P(x))] \\
&\quad + (1 - E_t p_{t+1}(x)) E_t [\mu(a_{t+1}^N)(1 - \delta_{t+1})(1 + r)u'(C_{t+1}^N(x))],
\end{aligned}$$

where $p_{t+1}(x)$ is defined as equation 6, and the second line comes from Envelop Theorem and Euler equations. Take the derivative of $E_t[V'_{t+1}(x)]$:

$$\begin{aligned}
\frac{1}{1+r} \frac{\partial E_t[V'_{t+1}(x)]}{\partial x} &= \underbrace{G'(\cdot) \left(E_t[(1 - \delta_{t+1})(\mu(a_{t+1}^P)u'(C_{t+1}^P) - \mu(a_{t+1}^N)u'(C_{t+1}^N))] \right)^2}_{> 0} \quad (7) \\
&\quad + \underbrace{p_{t+1} E_t[(1 - \delta_{t+1})\mu(a_{t+1}^P)u'' \frac{\partial C_{t+1}^P}{\partial x}]}_{< 0} + \underbrace{(1 - p_{t+1}) E_t[(1 - \delta_{t+1})\mu(a_{t+1}^N)u'' \frac{\partial C_{t+1}^N}{\partial x}]}_{< 0} \quad (8)
\end{aligned}$$

The second and third terms are negative because consumption is a normal good, and thus given the expected participation likelihood, if household has more capital, regardless of participation status, household would consume more in $t + 1$ and marginal expected present value would decrease. The first term, however, is positive. This is because the expected difference in marginal utility of contingent consumption (weighted by protection effect of participation, mathematically, $\mu(a_{t+1}^P)u'(C_{t+1}^P) - \mu(a_{t+1}^N)u'(C_{t+1}^N)$) also affects the expected

⁵⁴Derivation of $\frac{\partial E_t p_{t+1}}{\partial x}$ also depends on the difference between marginal utility of contingent consumption, weighted by protection of participation:

$$\frac{\partial E_t p_{t+1}}{\partial x} = (1 + r)G'(\cdot) E_t[(1 - \delta_{t+1})(\mu(a_{t+1}^P)u'(C_{t+1}^P) - \mu(a_{t+1}^N)u'(C_{t+1}^N))]$$

participation likelihood $E_t p_{t+1}$ in the same direction. Intuitively, when marginal utility of consumption when participating in $t + 1$ is higher (lower) from that of nonparticipating, household will have a higher (lower) likelihood of participation in $t+1$, which in turn increases the marginal expected present value. If the difference in marginal utility of contingent consumption is stark enough, the expected value function would exhibit nonconcavity for some range of capital value, in which case it is possible that $E_t[V'_{t+1}(K_{t+1}^P)] > E_t[V'_{t+1}(K_{t+1}^N)]$, thus not violating Euler equations.

Therefore, there is no clear prediction on the effect of participation on consumption or investment.⁵⁵ One possible scenario of a drastic difference in marginal utility of contingent consumption is that when participation brings a much higher (lower) protection effect—that is, $\mu(a_{t+1}^P) \gg (\ll)\mu(a_{t+1}^N)$. If there is no protection effect of participation detected, given proper assumption of utility function and density function of idiosyncratic intrinsic utility ζ_t , it is less likely to observe nonconcavity in the expected value function, and thus household would behave more normally (consumption increases in t when participating t).

Proof of Proposition 1. Rewrite the budget constraint for participation:

$$C_t^P + K_{t+1}^P = T + (1 + r)(1 - \delta_{it})K_t.$$

⁵⁵The logic on investment is similar. Suppose participants overconsume t so that they have lower capital level in $t + 1$, that is, $K_{t+1}^P \leq K_{t+1}^N$, $C_t^P \geq C_t^N + T - w$. With concave utility function, Euler equations and Mean Value Theorem,

$$\exists \theta \in [C_{t+1}^N, C_{t+1}^N + T - w], \quad \text{s.t. } E_t[V'_{t+1}(K_{t+1}^P)] - E_t[V'_{t+1}(K_{t+1}^N)] \leq u''(\theta)(T - w) < 0.$$

If the expected value function exhibits sufficient nonconcavity, the condition above is possible.

Take derivative with regard to A_k :

$$\frac{\partial C_t^P}{\partial A_k} + \frac{\partial K_{t+1}^P}{\partial A_k} \equiv (1+r) \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right]. \quad (9)$$

Take derivative of Euler equation 5 with regard to A_k :

$$u''(C_t^P) \frac{\partial C_t^P}{\partial A_k} \equiv H(K_{t+1}^P) \cdot \frac{\partial K_{t+1}^P}{\partial A_k}, \quad (10)$$

where $H(x) \equiv \frac{\partial E_t[V'_{t+1}(x)]}{\partial x}$. Solving equation 9 and 10:

$$\frac{\partial C_t^P}{\partial A_k} = \frac{1}{1 + u''(C_t^P)/H(K_{t+1}^P)} \cdot (1+r) \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right]. \quad (11)$$

Likewise, one can derive $\frac{\partial C_t^N}{\partial A_k}$ by simply replacing superscript P with N . To simplify the question, assume $u''(C_t^P)/H(K_{t+1}^P) = u''(C_t^N)/H(K_{t+1}^N) \equiv \gamma \in (0, +\infty)$.⁵⁶ Thus, the effect of attack on consumption, capital, and investment are as follows (superscripts are omitted because we assume participation does not influence the effect of attack on these variables):

$$\frac{\partial C_t}{\partial A_k} = \frac{1+r}{1+\gamma} \cdot \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right] \quad (12)$$

$$\frac{\partial K_{t+1}}{\partial A_k} = \frac{1+r}{1+1/\gamma} \cdot \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right] \quad (13)$$

$$\frac{\partial I_t}{\partial A_k} = \frac{r-1/\gamma}{1+1/\gamma} \cdot \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right] \quad (14)$$

The term $\left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta_t) \frac{\partial K_t}{\partial A_k} \right]$ is always negative because either attack increases current de-

⁵⁶Intuitively, γ measures the difference between concavity of consumption in t and that of continuation value in the future. $\gamma > 0$ is predicated on expected value function being concave (see the discussion above). If γ is higher, household's utility of consumption today reacts more to the exogenous shock than consumption tomorrow.

preciation rate ($\frac{\partial \delta_t}{\partial A_k} > 0$ if $k = t$), or past attack decreases past capital, which decreases current capital accumulation ($\frac{\partial K_t}{\partial A_k} < 0$ if $k < t$ by induction). Given the assumption $\gamma \in (0, 1)$, the effect of past attack on consumption and capital should be unambiguously negative. The effect on investment, however, depends on the comparison between the return to investment r , versus the concavity of value function $1/\gamma$, which determines the difference in marginal value between consuming today and in the future. \square

Proof of Proposition 2. Take differentiation of equation 6 with regard to A_k :

$$\begin{aligned} \frac{\partial p_t}{\partial A_k} &= G'(\cdot) \left[\frac{\partial \theta_t}{\partial A_k} + \frac{\partial (\mu(a_t^P) V_t^{P,s} - \mu(a_t^N) V_t^{N,s})}{\partial A_k} \right] \\ &= G'(\cdot) \left[\underbrace{\frac{\partial \theta_t}{\partial A_k}}_{\text{Intrinsic motivation}} + \underbrace{\left(\mu(a_t^P) \frac{\partial V_t^{P,s}}{\partial A_k} - \mu(a_t^N) \frac{\partial V_t^{N,s}}{\partial A_k} \right)}_{\text{Wealth effect}} + \underbrace{\left(\frac{\partial \mu(a_t^P)}{\partial A_k} V_t^{P,s} - \frac{\partial \mu(a_t^N)}{\partial A_k} V_t^{N,s} \right)}_{\text{Protection effect}} \right] \end{aligned}$$

where the second equation comes from Envelop Theorem and Euler equations. By assumption, intrinsic motivation would increase if household head experienced a past attack which leads to higher participation likelihood.

Wealth effect: Use Envelop Theorem and Euler equations, for $k \leq t$, the difference

between contingent present values of participation vs. nonparticipation.⁵⁷

$$\begin{aligned}
& \mu(a_t^P) \frac{\partial V_t^{P,s}}{\partial A_k} - \mu(a_t^N) \frac{\partial V_t^{N,s}}{\partial A_k} \\
&= (1+r) \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta) \frac{\partial K_t}{\partial A_k} - \frac{\partial C_t^P}{\partial A_k} \right] \mu(a_t^P) u'(C_t^P) \\
&\quad - (1+r) \left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta) \frac{\partial K_t}{\partial A_k} - \frac{\partial C_t^N}{\partial A_k} \right] \mu(a_t^N) u'(C_t^N).
\end{aligned}$$

Substitute equation 12 into the equation,

$$\begin{aligned}
& \mu(a_t^P) \frac{\partial V_t^{P,s}}{\partial A_k} - \mu(a_t^N) \frac{\partial V_t^{N,s}}{\partial A_k} = \\
& (1+r) \underbrace{\left(\frac{1}{1+1/\gamma} \right)}_{>0} \underbrace{\left[-\frac{\partial \delta_t}{\partial A_k} K_t + (1-\delta) \frac{\partial K_t}{\partial A_k} \right]}_{<0} \cdot (\mu(a_t^P) u'(C_t^P) - \mu(a_t^N) u'(C_t^N)). \quad (15)
\end{aligned}$$

Therefore, the wealth effect of attack depends on the term of $\mu(a_t^P) u'(C_t^P) - \mu(a_t^N) u'(C_t^N)$, i.e., difference in marginal utility of contingent consumption weighted by protection of participation. Intuitively, given the protection effect of participation, if household thinks participation brings higher consumption and thus lower marginal utility of consumption, it is more likely that he switches from nonparticipation to participation status, and vice versa.

In addition, wealth effect of attack on participation can be expressed in terms of the effect of attack on investment:

$$\mu(a_t^P) \frac{\partial V_t^{P,s}}{\partial A_k} - \mu(a_t^N) \frac{\partial V_t^{N,s}}{\partial A_k} = \frac{1+r}{r-1/\gamma} (\mu(a_t^P) u'(C_t^P) - \mu(a_t^N) u'(C_t^N)) \cdot \frac{\partial I_t}{\partial A_k}.$$

⁵⁷We assume here that attack does not affect labor income T or w . It is possible that past attack might affect respondents' productivity in different sectors, for instance, through handicap, and thus past attack might affect labor income. Our empirical analysis shows no additional effect of past attack that only affects respondents, and our main analysis focuses on attacks on other household members, which arguably does not affect the labor income of household head in principle. More discussion in Appendix F.

Therefore, if $\mu(a_t^P)u'(C_t^P) - \mu(a_t^N)u'(C_t^N)$ does not vary much, one can control for investment in time t to control for wealth effect of past attack on participation.

Protection effect: The direction is ambiguous. If expected function is well-behaved and concave, it is easy to show $V_t^{P,s} > V_t^{N,s}$ because participation given survival brings higher labor income and thus higher expected continuation value. If the subjective survival likelihood does not vary much with past attack, then protection effect is positive. If, however, the subjective survival likelihood decreases sufficiently when household experienced a past attack, household head would be less likely to participate. In the main analysis, we control for whether respondent overreports any violent event as a proxy to subjective survival likelihood to control for the protection effect.

□

F Attack on respondent himself

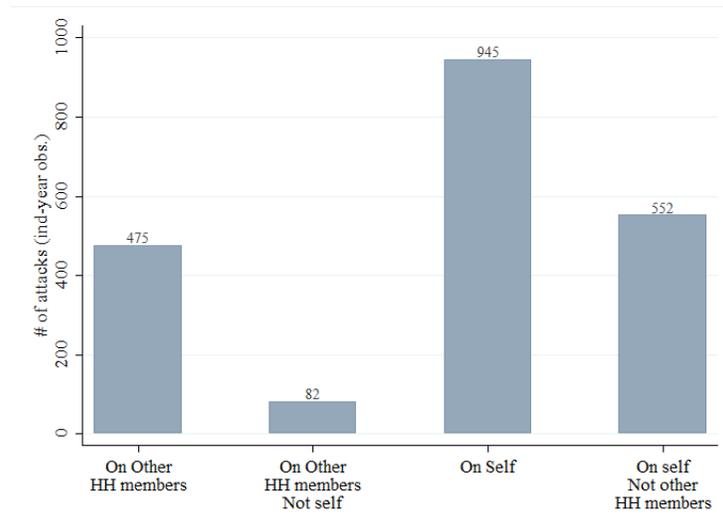
In the main text, we choose attacks on other household members as main attack definition instead of attacks on respondents themselves. The main concern of using attacks on respondents themselves, as we discussed in Section 8, is that attacks on respondents themselves can possibly affect respondents' physical capabilities of participation. The effect of past attack on respondents themselves is likely a mix of the negative effect from decreased physical capabilities and the positive effect from intrinsic preferences.

Figure F.1 shows the distribution of attacks on other household members and attacks on self. Out of 475 attacks where respondent's other household members are attacked, most of them involve respondent themselves being attacked; only 87 of these episodes do not involve respondent themselves being attacked. There are 945 episodes where respondents reported themselves attacked during the violent events; 552 of them do not involve other household members being attacked.

Table F.1 replicates the main results in Table II but breaks down the effect by attacks on other household members and on respondents themselves. We only examine attacks by foreign armed group and participation into a Congolese militia. Column (1) is the benchmark regression. Column (2) breaks down the main attack variable into attacks only involving other household members and attacks involving both other household members and respondents themselves. The effect of attack on both is statistically significant. The effect of attack only on other household members is less precisely estimated due to the insufficient number of such events (only 87), but the coefficient remains similar to the main result (2.92 vs. 2.55).

Column (3) examines the effect of attack on respondents themselves. The effect is a

Figure F.1: Distribution of Attack on Other HH Members and Self



Notes: This figure shows the distribution of attack observations where other household members or respondent himself was attacked.

precise zero. Column (4) breaks down the effect into attacks only involving respondents and attacks involving both other household members and respondents themselves. The effect of attack on both is still statistically significant, but the effect of attack involving only respondents is again a precise zero. Considering the large number of such events (552), we do not think the zero effect is due to imprecision. Column (5) includes attacks only involving other household members, attacks only involving respondents and attacks on both. The main conclusion remains unchanged. All the evidence suggest that attacks only on respondents cannot capture well the intrinsic preferences of participation because it potentially affects respondents' physical capabilities to join an armed group. We thus adopt attacks on other household members as the main definition of attack throughout the paper.

Table F.1: Main Regression, Separating Attacks on Other HH Members and Respondents

	Participation in Congolese militia in year t (%)				
	(1)	(2)	(3)	(4)	(5)
Past attack by foreign armed group					
On other HH members	2.55** (1.01)				
On other HH members, not on respondents		2.92 (2.43)			2.92 (2.43)
On respondents			-0.02 (0.64)		
On respondents, not on other HH members				-0.05 (0.73)	-0.02 (0.73)
On respondents and on other HH members		2.37** (1.05)		2.38** (1.05)	2.37** (1.05)
Control mean	1.89	1.89	1.82	1.82	1.81
Obs.	25060	25060	25060	25060	25060

Notes: This table presents OLS estimates of equation 1, with different attack definitions. The dependent variable is an indicator for whether the respondent joins a Congolese militia. The main explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household by foreign armed group before year t . P-value: *** 0.01, ** 0.05, * 0.10.

G Migration

G.1 Difference between migrants and stayers

We first look into whether villagers who migrated are systematically different. Table G.1 compares individual-year observations where villagers moved to a new village in year t versus those where villagers stay in the same village. In total there are 1,223 migration episodes. Migrants are more likely to have participated in any armed group before, but less in a Congolese militia. They tend to be younger, less likely to have married, less likely to have worked in mining or agriculture, and more likely to be in school or unemployed before moving. They do not differ in economic status before migration, although they invest more after they migrate. Migrants are not more likely to experience any attack by foreign armed group in the past, but are less likely to have experienced an attack by Congolese militia.

G.2 Migration as a source of selection bias

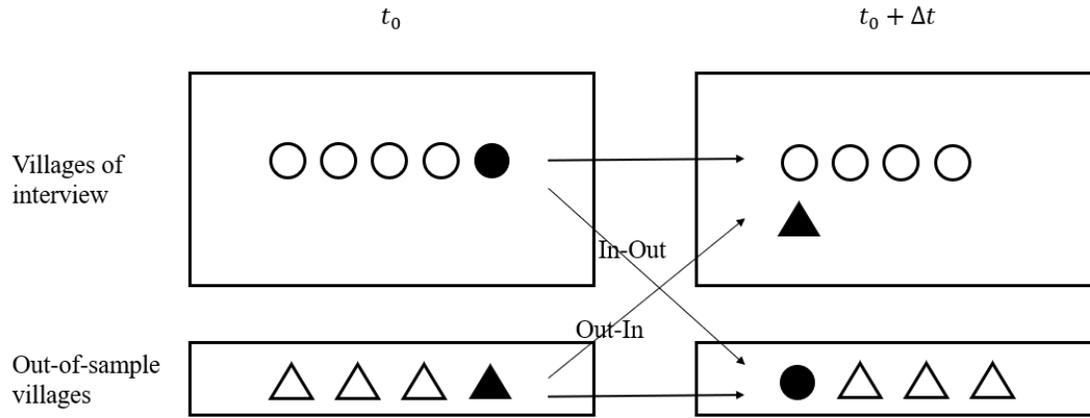
To discuss selection bias resulting from migration, it is necessary to first define the ideal average treatment effect (ATE) of the study. Suppose at time t_0 we have a representative sample from the villages we interview, and we want to estimate the treatment effect of past attack on villagers from the villages of interview. After a period Δt , however, some villagers emigrate to an out-of-sample village (In-Out migration), and some villagers migrate into a village of interview (Out-In migration). A random draw from the interview villages in $t_0 + \Delta t$ will not be representative of villagers from the villages we interview at time t_0 . Notice that some villagers migrate within villages of interview, but this does not cause the selection bias because they do not alter the composition of villagers from the villages of interview.

Table G.1: Characteristics of Migrants (1,223 migration episodes)

	Nonmigrant (1)	Migrant (2)
Obs. (Age \geq 15, 1995 \leq $t \leq$ 2013)	23,868	1,223
<i>Panel A: Participation history</i>		
Ever participated in any armed group before year t (%)	8.05	8.26*
Ever participated in Congolese militia in year t (%)	5.40	4.91***
Ever participated in foreign armed group in year t (%)	2.02	1.64
<i>Panel B: Socio-demographic background per individual-year obs</i>		
Age in year t	32.33	28.67***
Married in year t (%)	44.31	38.76**
Works primarily in mining in year $t - 1$ (%)	14.21	12.48**
Works primarily in agriculture in year $t - 1$ (%)	47.18	38.6**
Works primarily as civil servant in year $t - 1$ (%)	7.11	7.32
Works primarily in school in year $t - 1$ (%)	9.36	12.57**
Unemployed in year $t - 1$ (%)	22.14	28.99***
<i>Panel C: Economic status per individual-year obs</i>		
Wealth at birth (z-score, only $t = 2012$)	0.04	0.07
Asset stock in year $t - 1$ (z-score)	-0.03	-0.05
Investment in year $t + 1$ (z-score)	0.09	0.23***
<i>Panel D: Attack history per individual-year obs</i>		
Experienced attack on own HH by any armed group before year t (%)	12.08	11.45
Experienced attack on own HH by foreign armed group before year t (%)	9.32	9.08
Experienced attack on own HH by Congolese militia before year t (%)	3.02	2.13**

Notes: This table shows descriptive statistics for observations where respondents move to a new village versus those where respondents stay in year t . Economic indices are only computed in South Kivu, where the data are available. We indicate the difference between Column 1 and 2 (P-value: *** 0.01, ** 0.05, * 0.10), computed after including village FE and year FE, and clustered two-way at the individual respondent and the village*year level. Construction of economic indices (principal component analysis on following variables): Wealth at birth index—Stock of cows at birth, stock of goats at birth, stock of pigs at birth, stock of lands at birth, relation to village chief, number of father’s wives. Asset stock index—stock of cows, stock of goats, stock of pigs, stock of lands. Investment index—purchase of cows, purchase of goats, purchase of pigs, purchase of lands.

Figure G.1: Illustration of Selection Bias



More formally, see Figure G.1:

1. Suppose villages of interview (Group A) constitute proportion $a \in [0, 1]$ of the East Congo population.
2. Within villages of interview, proportion $1 - \pi$ of the villagers will never migrate outside (A_1). Proportion π of the villagers will migrate to out-of-sample villages at least once throughout the period (A_2) with probability p .
3. Within out-of-sample villages (Group B), proportion $1 - \pi$ of the villagers will never migrate outside (B_1). Proportion π of the villagers will migrate to villages of interview at least once throughout the period (B_2) with probability p .

Assume the real treatment effect of each group is $T(X)$, and past attack does not change

the composition of different subgroups. Treatment effect of interest:

$$ATE(A) = (1 - \pi)T(A_1) + \pi T(A_2).$$

Estimate of the treatment effect:

$$\begin{aligned} AT\hat{E}(A) &= \frac{a(1 - \pi)}{a(1 - \pi) + (1 - a)p\pi}T(A_1) + \frac{(1 - a)p\pi}{a(1 - \pi) + (1 - a)p\pi}T(B_2) \\ &= \frac{1}{1 + (\frac{1-a}{a}p - 1)\pi}ATE(A) + \frac{(1 - a)p\pi T(B_2) - a\pi T(A_2)}{a(1 - \pi) + (1 - a)p\pi}, \end{aligned} \quad (16)$$

There are at least two different ways that selection bias affects the estimation of ATE (through the coefficient term and constant). If, however, we assume that past attack does not affect the composition of migrants and nonmigrants, one can at least estimate the treatment effect on villagers who never migrate outside sample villages without bias ($T(A_1)$).

Table G.2 estimates the effect of past attack by Foreign armed group on participation in a Congolese militia. Column (1) shows the benchmark of the main result. Column (2) implements the main specification 1 within villagers who never migrate outside of the villages of interview. The effect on these villagers are larger and remain statistically significant. This is an unbiased estimate of the effect of past attack on villagers who never migrate outside of the villages of interview if past attack does not affect the composition of migrants ($T(A_1)$). Columns (3) and (4) implements the main specification but controls for different types of migrants. The estimates of $T(A_1)$ remain largely similar.

Results in Columns (3) and (4) also suggest that migrants from outside of the villages of interview, if anything, are negatively selected. Those who are less likely to participate tend

Table G.2: Effect of Past Attack on Migration, on Villagers Who Never Migrated Outside of Sample

	Participation in Congolese militia in year t (%)			
	(1)	(2)	(3)	(4)
Past attack on HH by foreign armed group	2.55** (1.01)	4.02*** (1.48)	4.12*** (1.44)	4.14*** (1.44)
Past attack * Migrants			-3.88** (1.75)	
Past attack * Out-In migrants				-2.76 (1.71)
Past attack * In-Out migrants				-5.86** (2.62)
Past attack * Yet-to-be migrants				-5.06*** (1.70)
Only within nonmigrants		✓		
Control mean	1.89	2.28	1.89	1.89
Obs.	25,060	12,710	25,060	25,060

Notes: This table presents OLS estimates of equation 1, controlling for migration. The dependent variable is an indicator for whether the respondent joins a Congolese militia. The main explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household by foreign armed group before year t . P-value: *** 0.01, ** 0.05, * 0.10.

to migrate into the villages of interview. Remember, however, that villagers who migrated from outside of the villages of interview are not representative of Out-In migrants or In-Out migrants, and thus these are potentially a biased estimate of $T(B_2)$ and $T(A_2)$.

G.3 Calibration of real ATE

We conduct a simplified calibration exercise using the results above, assuming migration does not affect migration likelihood (more discussion in the next subsection).

Migration likelihood for each migrant p : On average, each migrant is observed for 16 years in the sample, and moves on average twice. We calibrate $p = \frac{1}{8}$.

Proportion of villagers in the villages of interview a : According to village chief survey, on average, there are 427 villagers in a village of interview in South Kivu. Consider the total population in South Kivu in 2015 to be 5,772,000, and apply the average number of villagers to all 239 villages in South Kivu and North Kivu. We calibrate $a = 1.8\%$.

Proportion of villagers who migrate at least once throughout the observation period π : Out of 1,537 respondents, 1,086 have never migrated outside of the sample once. We calibrate $\pi = 1 - 1086/1537 = 29\%$.⁵⁸

Now we can apply these parameters to Equation 16. We have an imperfect estimate of $T(B_2)$ and $T(A_2)$ from Table G.2, Column 4 ($\hat{T}(B_2) = 4.14 - 2.76 = 1.38$, $\hat{T}(A_2) = 4.14 - 5.86 = -1.72$), and an estimate of $A\hat{T}E(A) = 2.55$.

Suppose all In-Out migrants do not react to past attack at all; that is, $T(A_2) = 0$. To make the real $ATE(A)$ zero, one needs the effect on Out-In migrants $T(B_2)$ to be at least

⁵⁸This is not the perfect calibration because the denominator does not consider villagers who have emigrated. It is hard to observe emigrants who left the sample; the closest data we have is how many villagers emigrated from each village every year. On average, a village of interview in South Kivu sees 61 in-migrants and 64 out-migrants every year.

3.46. This is different from $\hat{T}(B_2) = 1.38$ with statistical significance (p-value 0.0205). To justify the empirical result, one needs to assume that the representative migrant from outside of the sample reacts more strongly to the past attack, but we only observe the left tail of the distribution.

Suppose all Out-In migrants do not react to past attack at all; that is, $T(B_2) = 0$. To make the real $ATE(A)$ zero, one needs the effect on In-Out migrants $T(A_2)$ to be at least -23.63. This is drastically different from $\hat{T}(A_2) = -1.72$ (p-value 0.0000). To justify the empirical result, villagers who tend to migrate outside of the villages of interview should be strongly discouraged from participation if they have experienced any past attack by foreign group, but we only observe the extremely right tail of the distribution.

G.4 Migration as an alternative channel

Selection bias aside, migration can confound the main result in the following way. Attack history in the past might affect the propensity of migration, and migration leads to a different participation rate through another channel that we did not discuss in the main text.

Table G.3 regresses different migration indicators on whether respondent experienced an attack by foreign armed group before. Column (1) suggests that a past attack leads to lower propensity of migration, although the effect is mainly driven by migration within the villages of interview (see Column (4)), which does not trigger the selection bias discussed above. Column (5) and (6) control for migration history in the past. Although migration within the villages of interview seems to be positively correlated with higher participation rate in a Congolese militia, the main effect of past attack remains largely unaffected, which

suggests that migration does not explain the effect of past attack on participation.

Table G.3: Migration as a Potential Channel on Participation

	Migration (%)				Participation (%)	
	All (1)	In-Out (2)	Out-In (3)	In-In (4)	Militia (5)	Militia (6)
Past attack on HH by foreign armed group	-1.50*	0.10	-0.50	-1.02*	2.53**	2.41**
	(0.89)	(0.28)	(0.59)	(0.56)	(1.00)	(1.00)
Ever migrated before					1.54*	
					(0.81)	
Ever migrate from Out to In						0.14
						(1.68)
Ever migrate from In to Out						-0.28
						(1.00)
Ever migrate between villages of interview						4.74***
						(1.30)
Control mean	4.80	0.81	2.03	1.45	1.42	1.40
Obs.	25,060	25,060	25,060	25,060	25,060	25,060

Notes: This table presents OLS estimates of equation 1. Column (1) to (4) use different migration indicators as the dependent variable (migration in general, migration from a village of interview to an out-of-sample village, migration from an out-of-sample village to a village of interview, migration between the villages of interview). In Column (5) and (6), the dependent variable is an indicator for whether the respondent joins a Congolese militia. The main explanatory variable is an indicator for whether the respondent's household has been attacked by foreign armed groups. Column (5) and (6) control for whether respondent has migrated in the past. We include observations between 1995 and 2013 above age 15 at year t . All regressions include individual FE, village FE, year FE, age FE, and cluster two-way at the individual respondent and the village*year levels. Control mean is computed among observations where respondents never experienced an attack on household by foreign armed group before year t . P-value: *** 0.01, ** 0.05, * 0.10.