

# Rebel Group Formation in Africa: Evidence from a New Dataset

Forthcoming at *World Development*

Janet I. Lewis<sup>1</sup>  
June 2022

Theories of internal armed conflict onset often conceptualize an eruption emerging from pre-existing, non-violent opposition movements or organizations. This paper argues that this is not the most common pathway to armed conflict in Africa. It introduces new data on rebel group formation in 47 African states (1997-2015) that offers a more complete picture than existing conflict datasets; about two-thirds of the groups in the new dataset are not named in standard conflict datasets. The strong majority of these rebel groups formed in rural areas; did not emerge from nearby protest movements or civil society organizations; and initially did not benefit from substantial material resources nor commit large-scale violence. Analyses suggest that there are more instances of new rebel group formation in states with more rural territory – which I interpret as indicating lower state monitoring capacity, and thus lower barriers to entry for new rebel groups – and that rebel groups forming in rural parts of African states draw less often on pre-war, non-violent contestation than those of urban areas. The paper highlights the need for more theory on the varieties of pathways to internal armed conflict, and the importance of – and challenges of collecting – systematic evidence on armed group formation.

Internal armed conflict; civil war; insurgency; rebellion; protest; Africa

---

<sup>1</sup> Associate Professor of Political Science, George Washington University, janetilewis@gwu.edu. For excellent research assistance, I thank Jennifer Hamilton and Andy Marshall, as well as Ebunoluwa Akinbode, Sohini Ashoke, Mark Thomas-Patterson, Joseph Schluger, Ethan Shuchart, Andy Tomusiak, and Kaitlyn Yuan. The paper benefited from helpful comments from Jessica Braithwaite, Jennifer Hamilton, Iris Malone, Anne Meng, Noah Nathan, Amy Poteete, Stephen Rangazas, and Rob Williams. I gratefully acknowledge funding from the Minerva R-DEF program and George Washington University to support data collection. All mistakes are my own.

## I. Introduction

Why and how does internal armed conflict start? Most of the large existing literature on the onset of civil conflict – armed conflict between a state and at least one organized, non-state actor – tends to approach this question with the aim of understanding a single process that governs it. Existing research tends to distinguish the causes of organized, internal armed conflict from that of disorganized violence like riots or intercommunal violence (e.g. D. Cunningham and Lemke 2014) and distinguish the drivers of violent versus non-violent mobilization (Chenoweth and Stephan 2013; K. Cunningham 2013; D. E. Cunningham et al. 2017; Thurber 2021). Yet the possibility of multiple, distinct pathways to civil conflict onset is rarely considered.

There is good reason to question whether similar processes drive, for example, both the emergence of rebel groups in Syria in late 2011 – which occurred after months of urban, public protest and military implosion (Mazur 2021) – and the numerous instances of localized rebel group formation in remote eastern Democratic Republic of Congo (DRC) in roughly the same period. Even within a single civil war, there is variation in different anti-state armed groups' early stages. For example, in the civil war in South Sudan, the largest rebel group, the SPLA/M/IO, formed in the capital, Juba, drew heavily on its namesake rebel organization-turned-ruling party from which it emerged and challenged. In contrast, the smaller TFNF started in rural Upper West Nile State, far from the capital, and drew to a lesser degree on prior organizations.<sup>2</sup>

This paper proposes the importance of considering variation in pathways to anti-state armed group emergence – what I call rebel group formation – especially in regard to when and

---

<sup>2</sup> These groups are included in the dataset on which this paper relies; sources and details on these groups can be found in the detailed “Record of Coding Decision” document available alongside the dataset.

how political entrepreneurs in rural versus urban areas mobilize local populations and build armed organizations. The paper's primary aim is to demonstrate the promise of rebel group formation as an outcome of interest, and focus of data collection efforts, by introducing and exploring new data on rebel group formation in Africa.

Building on several earlier theories of civil conflict onset, insurgency, and rebel group emergence (e.g. [Leites and Wolf 1970, Chapter 7](#); [Fearon and Laitin 2003](#); [Lewis 2020](#)), I propose that African rebel groups are more likely to form – that is, to begin operational planning for, and conducting, violent operations against the state – in rural areas, where it is more difficult for states to gain information needed to detect and defeat them. Additionally, I argue that groups forming in rural areas are less likely to form out of recent, public episodes of local mobilization than those forming in urban areas; historically, states have had less control and less internal monitoring capacity beyond their urban core ([O'Donnell 1993, 1358](#); [Scott 1998, 181–92](#)), and aspiring rebels face lower barriers to entry where the state is barely present ([Weinstein 2007, 14, 341](#)). Furthermore, in low-income African countries, strong civil society organizations that lend themselves to anti-state mobilization are often scarce in rural areas. While organizational endowments are useful if they are available, in rural areas where the state's monitoring capacity is relatively weak, aspiring rebels do not *need* to rely on extensive, pre-mobilized networks in order to begin their violent campaigns. Instead, nascent rebels can begin as a small group, committing only small-scale violence for a prolonged period as they build and grow their organization. In urban territories where states tend to have higher monitoring capacity, armed group formation does sometimes occur<sup>3</sup> – but nascent rebels will be more quickly discovered, located, and defeated if they do not have an already-formed

---

<sup>3</sup> See especially [de la Calle and Sánchez-Cuenca \(2012\)](#) for a related argument that *territorial* rebels are most likely form in the weakest states, whereas *non-territorial* armed groups – which are underground groups that principally engage in urban terrorism – are most likely to form in stronger state contexts.

organization or pre-mobilized group from which they can draw on to clandestinely plan and quickly build organizational capacity for anti-state violence.

Much of this argument echoes prior theories but has not yet been demonstrated with systematic evidence on rebel group formation nor integrated into dominant theories of internal armed conflict onset. Data limitations have been the primary obstacle. There are major challenges to observing the full range of instances of anti-state armed group emergence – especially in the period immediately following a group’s decision to violently challenge a state, as they begin planning and conducting their initial attacks. These early, often clandestine phases of group formation are especially difficult to observe in remote areas, where the state and news media’s presence is lowest. Yet examining this variation is vitally important to understanding the precursors of armed conflict.

This paper brings new data<sup>4</sup> to bear on the early phases of armed conflict in African states – about 65% of which have experienced substantial armed conflict since independence (Straus 2012).<sup>5</sup> The new data introduced here captures a much larger portion of rebel group formation in Africa from 1997-2015 than standard conflict datasets; about *two-thirds* of the rebel groups in this data are not named in the Uppsala Conflict Data Programme/Peace Research Institute of Oslo Armed Conflict Dataset (UCDP/PRIO ACD) (Gleditsch et al. 2002) or the Global Terrorism Database (GTD).<sup>6</sup> This is especially noteworthy since this new data

---

<sup>4</sup> The dataset, codebook, and a lengthy (more than 700 page) document indicating rationales behind, and sources used, to decide which groups to include and for coding variables, and all groups named in ACLED but excluded from this dataset, is available at the Harvard Dataverse [url provided upon publication].

<sup>5</sup> To arrive at this figure, Straus uses the conventional definition of armed conflict as an instance of organized, anti-state violence reaching at least 25 recorded, battle-related deaths.

<sup>6</sup> This is consistent with other recent efforts to use event data in conjunction with extensive additional research to count instances of armed group formation. See especially Malone (Forthcoming), which uses the population of violent non-state organizations listed in Asal et. al. (2015) that conducted at least one violent attack in recorded in the Global Terrorism Database (GTD) to arrive at a new dataset of 1,202 armed groups globally (1970-2012); many more groups than prior efforts.

intentionally omits instances of rebel splinters and mergers, aiming to capture only new rebel groups.<sup>7</sup>

While the new data is coarse due to a lack of available detailed evidence – all variables are dichotomous, with considerable missing data on most variables – it reveals new patterns of rebel group formation across Africa since the late 1990s. In particular, it shows that rebel group formation is much more common than standard datasets indicate, and that it often starts in a manner different from that of an violently-erupting, widespread movement: The modal rebel group appears to have started in a rural area, lacked substantial material endowments, did not follow in the wake of a local, nonviolent mobilization like a protest, and initially engaged in only small-scale violence (no incident resulting in more than 25 battle deaths in the first three years). Consistent with the paper’s theoretical arguments, basic analyses suggest that countries with more rural territory have more instances of rebel group formation, and that rebel groups that form in rural areas are less likely than those forming in urban areas to form in the aftermath of a local protest, or to rely heavily on prior, nonviolent organizational endowments.

This paper makes several contributions to the study of armed conflict in developing countries. First, it demonstrates the need for a wider aperture in our concepts and evidence about the landscape of organized armed rebellions’ start. Specifically, it indicates the promise of rebel group formation as an outcome of interest, and of considering how different factors – such as urban versus rural contexts – condition a decision of whether and how to rebel violently against the state. Second, the paper contributes to a burgeoning literature using new data to examine the relationship between non-violent and violent mobilization (e.g. (Ryckman 2020;

---

<sup>7</sup> We define a splinter group as occurring when *most* leaders and *most* members of the new group come from the same, pre-existing “parent” rebel group. If we found evidence that just a small portion of the initial leaders and members of a new rebel group come from a parent group, we do not count this as a splinter but instead as a “new” rebel group, and include it in the dataset.

Braithwaite and Licht 2020; Butcher et al. 2021). It adds descriptive evidence that in Africa, organized violence against the state often does *not* emerge out of non-violent mobilization such as protests – even though rebel groups may hope to instigate urban protests (Leventoğlu and Metternich 2018) and articulate agendas that resonate with urban populations (Mkandawire 2002).

Third, the paper demonstrates both the potential and the challenges of collecting systematic data on armed group formation. It joins other new efforts (see especially Malone Forthcoming) in demonstrating that it is possible to do so for a wider array of groups than are currently included in standard datasets – by capturing groups committing only small-scale violence, and capturing them from an earlier stage of inception – using publicly-available information. This indicates the promise of continued efforts to collect better data about initial rebel group formation in a wide range of contexts. Doing so may also help to reconcile divergent findings among prior research at different levels of analysis – and microlevel studies from different states – about the processes of mobilization during conflict onset and escalation (Balcells and Justino 2014). On the other hand, this paper demonstrates the difficulty of gaining complete, fine-grained information about all armed groups forming at a given place and time. For example, as I describe below, the data collection exercise that underpins this paper indicated that numerous rebel groups likely existed, but insufficient information was available to code any detail about them. Also, it was not possible to capture high-resolution information about the location of several groups' formation, which is critical for fine-grained subnational analysis. These limitations point to the importance of future data collection efforts gaining information from the localities where rebel groups reportedly formed to supplement what is available on the internet and in libraries. I return to these points and note promising initiatives in the conclusion.

## II. Starting Organized Violence With, and Without, Prior Mobilization

Several theories of civil conflict onset represent it with straightforward collective action dynamics like that of a protest. By these accounts, rebel group formation and violence occur *after* a period of public, non-violent mobilization of the local population. For example, classic works on revolutions posited that “lightning mobilization” of a “revolutionary mass” with widespread claims against a state would precede a more violent stage of conflict (Tilly 1978, 201; see also Moore 1984). More recently, prominent work theorizes a causal pathway to civil conflict onset that starts with group-level inequality, which leads to widespread grievances among groups, which in turn leads to frustration, mobilization, and then ultimately an increased risk of internal violent conflict (e.g. Cederman, Gleditsch, and Buhaug (2013)). New data has also facilitated advances in analyses of the organizational antecedents to civil conflict (Braithwaite and Cunningham 2020), showing that rebels’ “parent” organizations range from militaries, other rebel organizations or armed groups, political parties, trade unions, student movements, and others.<sup>8</sup>

While the specific mechanisms linking mobilization and organizations to rebel group formation vary in such accounts, a common assumption is that the more people in a given area who are pre-mobilized and organized, then the more likely it is that conflict will erupt into sudden, substantial violence. Such non-violent, prior mobilization is assumed to facilitate recruitment into the rebel army, and willingness among local civilians to collaborate. Indeed, prior movements or organizations – violent or even nonviolent ones – can allow aspiring rebels

---

<sup>8</sup> The new Foundations of Rebel Group Emergence (FORGE) dataset, introduced in Braithwaite and Cunningham (2020), codes the organizational basis of the rebel groups on the UCDP/PRIO Non-State Actor list.

to quickly develop systems of command and control, use established networks of potential resources, and draw on bonds of trust born of prior, regularized contact (Zukerman-Daly 2012; Sarah Elizabeth Parkinson 2013; Staniland 2014). Without such networks and resources, incipient rebel groups need to spend time – while evading government authorities – acquiring the people, resources, and systems of command and control needed to carry out medium or large-scale attacks against more powerful, conventional state military forces (Parkinson and Zaks 2018).

On the other hand, James Fearon and David Laitin’s seminal article argues against the need for pre-mobilized groups in order to start a rebellion. They posit that insurgency does not require substantial military capacity, and thus only a handful of aggrieved individuals are needed to start one (Fearon and Laitin 2003, 76). A logical extension of this argument is that rebel groups can form and initiate violence irrespective of whether the local population has been widely mobilized and organized, or even feels strongly about the rebels’ cause. Evidence on internal warfare in Africa tends to support this latter description, finding that it is most commonly small-scale, with weak rebels that are “poorly structured and trained” and operate on states’ peripheries (Straus 2012, 181). This is especially true at the start of an armed campaign. For example, consider the National Resistance Army (NRA) rebels, who overtook the Ugandan government in January 1986 and have since governed as the ruling National Resistance Movement (NRM) party. While they are remembered for their high levels of social endowments as a disciplined rebel organization with strong ties to the local population (e.g. Weinstein 2007, 71), they started with just 26 men and one gun in western Uganda, far from the capital. In his memoir, NRA rebel leader and now (since 1986) Ugandan President Yoweri Museveni describes that initially, his force was not yet ready for large, risky confrontations and that their first attack was one aimed at the modest goals of capturing weapons and “helping our



soldiers overcome fear and see for themselves that it was feasible to destroy our opponents and still preserve our still very small force” (Museveni 1997, 132). The NRA built its organization and ties to local communities slowly, and at first clandestinely, as they conducted small acts of violence (Kasfir 2005, 279–80). Theory on rebel group formation and detailed evidence from Uganda also supports this argument, showing that incipient groups there looked much more like the small, secretive groups that Fearon and Laitin envisioned than the other approach sketched above (Lewis 2017; 2020). Those analyses also indicated that in Uganda, widespread grievances, or material or other endowments, did not drive the initial stages of rebel group formation. Similarly, Leventoğlu and Metternich (2018) demonstrate that armed conflicts in Africa typically begin in rural contexts – but the groups that manage to grow stronger, subsequently triggering urban protest, are more likely to gain concessions from the government.

These accounts from Africa starkly contrast with, for example, the start of insurgent violence in urban Karachi in the 1980s, which occurred after years of mobilization and organization of student activist groups (Staniland 2010). Similarly, the start of Sendero Luminoso’s violence in Peru occurred after a decade of clandestine mobilization and organization-building.<sup>9</sup> What accounts for the different sequences of mobilization and violence we see in these theories and examples of internal armed conflict onset? I propose below that the first characterization of rebel group formation sketched above – that of rebels emerging out of, and benefitting from, pre-mobilized, non-violent political movements or organizations – is more common in contexts where states have higher capacity to *monitor their territory*. In such contexts, nascent rebels simply need these endowments more in order to have a chance to survive.

---

<sup>9</sup> See especially Blaxland (2020) for in-depth discussion of how Sendero’s lengthy “incubation” during the 1970s led to an enduring insurgency throughout the 1980s.

Variation in this capacity occurs both across states, and within states; some states have higher monitoring capacity in general, and, as I substantiate below, many states are better at it in their urban areas than in their outlying rural territories. This is due to the greater ease of controlling and policing populations in cities (Kalyvas 2006, 133; Kocher 2004, 24), and the lesser average availability of collective action capabilities in rural areas, especially in Africa. I develop these arguments below.

Where the first characterization is supported by findings based on global conflict datasets, this may be due in part to selection problems in the dataset. As the authors of the codebook for the UCDP/PRIO Armed Conflict Dataset explain, “The bias produced by [the dataset’s] approach [of using news media accounts] is against the inclusion of conflicts in the earlier decades *and in the less-developed world*” (emphasis added, 3).<sup>10</sup> These problems may be particularly severe if rebel groups in weaker state contexts tend to be clandestine and not particularly violent when they initially form – a pattern I show below to be common in Africa. Media coverage of conflict events in Africa has also been shown to be much poorer than the rest of the world (Dietrich and Eck 2020),<sup>11</sup> and to suffer from urban bias in major conflict settings like South Sudan (Dawkins 2020, 13). This suggests that the magnitude of African rebel group omissions in standard conflict datasets may be quite large, and this problem is probably most

---

<sup>10</sup> Codebook version 18.1, available at [https://ucdp.uu.se/downloads/replication\\_data/2018\\_c\\_666956-1-1-k\\_ucdp-prio-acd-181-codebook.pdf](https://ucdp.uu.se/downloads/replication_data/2018_c_666956-1-1-k_ucdp-prio-acd-181-codebook.pdf)

<sup>11</sup> Dietrich and Eck present suggestive evidence in their article appendix that the lack of media reporting on conflict events in Africa relative to other continents is due to a general lack of news reporting there relative to other continents: The average number of news articles in Factiva on “politics/international relations” for the 17 conflict-affected countries in 2018 in Africa was 1,008 whereas the average was 9,065 articles for 16 non-African, conflict-affected countries.

severe in rural contexts. Yet these datasets are the evidence base upon which the strong majority of recent research on internal conflict onset relies.<sup>12</sup>

### **III. Argument**

#### *Weak State Monitoring Capacity in Rural Territories and the Frequency of Rebel Group Formation*

To begin shedding light on the processes of armed group formation in Africa, I return to a core question in the civil conflict onset literature: Why does rebel group formation occur more frequently in some states than others? I define rebel group formation as occurring when people come together to form an organization – adopting a command-and-control structure known and adhered to by group members – that aims to challenge the political authority and territorial control of the state using violence, making concrete plans to commit at least one act of violence against the state. Note from above that this group may be quite small, and I make no assumptions about the motives of the initial rebel leaders beyond their desire to at least minimally control territory. If we accept that all nascent rebels want to, at a minimum, survive to become viable groups, then regardless of the nature and intensity of a group’s motives, they are comparable during the formative period examined here. I posit that this is the case since groups must be minimally viable in order to extract concessions from the government, or become attractive allies for other rebels with whom they could join and do so, or even predate upon local resources or individuals.

---

<sup>12</sup> Among all quantitative articles on armed conflict onset published in ten top political science or conflict journals from 2003 through 2019, 78% rely on one of just four standard conflict datasets, of which the UCDP/PRIIO data is the most fine-grained, in that it uses the lowest threshold of recorded violence – 25 battle related deaths in one calendar year – for inclusion (Lewis 2020, 5–6).

Therefore, secrecy is what rebels need most at the very outset of their group formation; with just basic information about the rebels' identity and location, the government could end nascent rebels in this early, vulnerable stage. In making this assertion, I join a body of scholarship on insurgency and counterinsurgency that stress the fundamental importance of secrecy to the survival of insurgents (e.g. Fearon and Laitin 2003; Kalyvas 2006; Kilcullen 2010; Haim 2019), especially in the initial stages of insurgent group formation (Malone 2019; Lewis 2020).

I argue that states' capacity to monitor threats emerging from its territory is important to would-be rebels' calculation about whether or not to form a group, and that this capacity is typically much weaker in outlying, rural territories. Even weak states are almost always militarily stronger than incipient insurgents. Therefore, their capacity to detect and learn about them is their primary challenge in the early stages of counterinsurgency. Strategic, would-be rebels know this, and thus higher state monitoring capacity in a given area is key to deterring rebel groups from forming there in the first place. In linking rural territory to state monitoring capacity, I draw on well-established scholarship establishing African states' persistently limited penetration of outlying, rural areas (Coleman 1977; Herbst 2000), due largely to the high cost of doing so and greater strategic and economic importance of urban centers.<sup>13</sup>

While urban insurgencies do sometimes occur (Staniland 2010; Koren and Sarbahi 2018), urban areas are typically less hospitable to insurgents than rural areas (Kalyvas 2006, Kocher 2004). It is easier for states to monitor cities because of their relatively smaller territories and

---

<sup>13</sup> These dynamics may be different outside of Africa due a lower historical prevalence of small holder farming, and an earlier informational penetration of the state into rural areas. For example, in *Seeing Like a State*, James Scott links Southeast Asian colonial and post-colonial states' desire for monitoring their rural areas to their imposing plantations and other centralized agricultural schemes, despite those schemes' production inefficiencies. In fact, Scott refers to one land management scheme in independent Malaysia as "soft' civilian versions of the new villages created as part of counterinsurgency policy" (Scott 1998, 198-191).

higher population density, and because greater social heterogeneity and higher average incomes (and thus higher opportunity costs of rebelling) in urban areas make them ripe for anti-insurgent informants. In their classic article, Fearon and Laitin (2003, 80) explicitly link the drawbacks of urban insurgency to the state's informational advantages there, arguing, "in the city, anonymous denunciation is much easier to get away with, giving the government an advantage in its counterinsurgent efforts." Such assertions are supported by practitioners of counterinsurgency and insurgency. In the words of one U.S. Department of Defense study, "Usually, it is a mistake for the guerrilla to move into the city. In the city, the guerrilla is surrounded by a thousand eyes and a thousand jealousies" (Elkhamri et al. 2005, 1). Similarly, guerilla leader Fidel Castro stated, "the city is the grave of the guerilla" (de la Calle and Sánchez-Cuenca 2012, 581).

Two straightforward hypotheses about rebellion emerge from these theoretical foundations: *Rebel groups are more likely to form in rural areas than urban areas (H1a)*, and *new rebel groups are more likely to emerge in states with more rural territory (H1b)*. These hypotheses are not particularly novel, but they are rarely examined with systematic data on rebel group formation, as this paper aims to do.

#### *Prior Mobilization, Organizational Endowments, and Rebel Group Formation*

Conflict scholars have established the central importance of organizations in producing anti-state violence (e.g. Weinstein 2007; Sinno 2008; Staniland 2014; Balcells 2017; Parkinson and Zaks 2018). One implication of the discussion above is that while pre-existing organizational endowments – such as structures of hierarchy, roles, and procedures, or even material endowments – would be helpful for aspiring rebel leaders in any context, in rural contexts where state monitoring capacity is weak, would-be rebel entrepreneurs may not initially need

them to form a rebel group. Instead, they are more likely to calculate that in such a context, they could remain a small group, managing to maintain secrecy long enough to build an organization from the ground up, screening new recruits slowly and carefully as they begin committing violent attacks. In such contexts, rebel entrepreneurs can make gains with only modest increases in their coercive capacity. If they build a viable group, they can merge with other rebel groups, attract the attention of an external sponsor, or instigate urban protests, any of which would substantially increase their bargaining position with the central government. In contrast, where states have better informational penetration of their peripheral territories – and other actors have greater certainty that states will learn about and end insurgencies early – barriers to entry for armed groups challenging the state are higher. In these contexts, would-be weaker rebellions will more likely be disinclined to form without organizational endowments that could allow them to quickly mobilize and build a relatively large, capable armed group.<sup>14</sup>

A second reason why prior non-violent organizations and protest mobilization will be less likely to be important precursors to armed group formation in rural areas is that the types of organizational life and level of politicization needed to quickly launch a rebellion are rarer there. As Kalyvas (2007, 43-46) documents, the urban bias prevalent in conflict studies tends to privilege “top-down” perspectives that emphasize “high” politics and ideological or normative motivations for warfare, whereas rural civilians often have shorter time horizons and focus on more immediate social dynamics. Subsistence farmers’ participation in associational life tends to focus on *local* economic, social, and political issues.

---

<sup>14</sup> See for example, Petersen (2001, Chapter 3) for a discussion of how decades of rural reading circles and other clandestine opposition activities facilitated rapid mobilization for high-risk resistance and rebellion in Lithuania in the early 1940s. See also Malone (2019), which argues that in the rarer instances of armed groups forming in strong states, they can become viable due to miscalculations, shocks to state capacity, or selection effects.

Turning to Africa, scholarship on rural life there supports these points. For example, in his seminal work on the political economy of Africa, Robert Bates uses evidence from Nigeria, Zambia, Kenya, and Ivory Coast to show that the “numerous and widely scattered” nature of rural, smallholder farmer puts them at a substantial disadvantage in collective action (Bates 1983, 87–90). He goes on to note the “neutralization,” and “isolation” of rural farmers, and that “the very policies that oppress the peasantry... promote [their] political disorganization” (Bates 1983, 131–32). Thandika Mkandawire argues that, in contrast to the organized confrontations between peasants and landlords that have characterized agrarian relations in Latin America and Asia, the smallholder nature of farming in rural Africa has led farmers to adopt more “everyday,” evasive strategies of resisting the state (Mkandawire 2002, 199, quoting Scott 1985). Others have noted that civil society in African states tends to be small-scale and locally-focused (Bratton 1989, 410–411), or too weakened – often intentionally by governments beholden to urban interests – to exert substantial, oppositional force on politics (Gyimah-Boadi 1997, 280), at least in part due to patronage politics (Kasfir 1998, 126). Similarly, at the turn of the century, Jeffrey Herbst observed the lack of political parties with roots in rural Africa and stated, “It still appears too difficult to organize the peasants qua peasants” (Herbst 2000, 20).

Of course, vibrant civil society organizations from trade organizations to non-governmental advocacy organizations do thrive in African states, but they typically predominate in mining towns or more industrialized areas (Bratton 1989, 427).<sup>15</sup> Similarly, while anti-state protests have occurred in Africa with regularity since the colonial period, and increasingly so as the continent urbanizes<sup>16</sup> (e.g. Golooba-Mutebi and Sjögren 2017), this

---

<sup>15</sup> Consistent with this characterization, Brass (2016, 111) finds that number of NGOs in districts of Kenya are positively related to population density, and negatively related to distance from Nairobi.

<sup>16</sup> Rapid urbanization in Sub-Saharan Africa today is changing organizational life there; however, even as of 2018, about 60% Sub-Saharan Africa’s population was rural. Almost three quarters was rural in 1990. In

mobilization has remained a predominantly urban phenomenon (Branch and Mampilly 2015; Mueller 2018). Cross-national evidence from Africa shows that the percent of African countries' GDP comprised of manufacturing and the percent of the population that is urban are both positively associated with the onset of non-violent, anti-government campaigns; the former is also negatively associated with civil conflict onset (Butcher and Svensson 2016).<sup>17</sup> While civic life in rural Africa is of course complex and varied,<sup>18</sup> oppositional, nationally-oriented interest groups that underpin social movements and civic organizations have historically been less important to daily life there. Rather, their presence is more germane to contentious politics in Africa's cities.

In sum, I propose that because of would-be rebels' desire to avoid detection by the state – especially during the most vulnerable, nascent phase of insurgency – and the relative lack of strong oppositional organizations in rural Africa, *rebel groups forming in rural areas are less likely to emerge from recent episodes of mobilization such as protests (H2a) or civil society organizations (H2b) than those forming in urban areas.* While starting a rebellion with a pre-mobilized group or with substantial organizational resources would be preferable, aspiring rebels do not need them to start a rebellion in environments where the likelihood is low that the state will quickly meet them with military or intelligence pressure.

These arguments do not deny the usefulness of prior endowments to rebels in rural areas if they do exist, nor their importance to later stages of armed conflict. For example, after a rebel group grows into a substantial organization and violence escalates, rebels need – and are better able to absorb – larger quantities of recruits and material resources. At that stage, people

---

contrast, Latin America's population was about 20% rural in 2018 and 30% rural in 1990; South Asia's population was 66% and 75% rural, respectively. Data comes from the World Bank Data Bank based on United Nations Population Division's World Urbanization Prospects data.

<sup>17</sup> Cunningham et al. 2017 find that civil society organizations are negatively associated civil conflict onset, and positively associated with the onset of non-violent campaigns.

<sup>18</sup> See especially Boone (2003).



or practices from a pre-existing social movement or organization may be crucial. Also, once rebels are strong enough to govern residents of a given territory, pre-war community structures or other attributes would likely be relevant, as Arjona (2016) shows. These are later stages of conflict than those at the center of this paper.<sup>19</sup> The key point is simply that these are not as likely to be an important precursor to rebel group formation in rural areas, because aspiring rebels do not need them to get started. This has important implications for how we conceptualize the early phases of conflict; for example, indicating the importance of understanding how nascent rebels use violence and other means to mobilize populations that were not previously well-mobilized. I turn now to introducing a new dataset from Africa on rebel group formation and then probing the evidence for these arguments by examining descriptive patterns in the dataset.

#### **IV. Dataset on Rebel Group Formation in Africa**

With a team of research assistants, I compiled a new dataset that aims to include all non-state armed groups that formed between January 1, 1997, and December 31, 2015, that had political goals<sup>20</sup> and made concrete plans to violently challenge the authority of a state on the African continent (47 countries).<sup>21</sup> About 60% of these African states had at least one rebel group form and target them during this period; the number of rebel groups targeting each state we included in the dataset is shown in Figure 1 below. We did not include African island

---

<sup>19</sup> Aspiring rebels may anticipate these later stages, but I posit that their immediate concern is forming where they know the social and geographic terrain well, so their options for initial locations are limited.

<sup>20</sup> Because rebel groups often have strategic incentives to avoid initially publicly articulate political goals, we inferred political goals if the target(s) of initial violence appeared to be political in nature (especially targeting agents of the state), or if the nascent armed group sought to control territory.

<sup>21</sup> Analyses in [citation removed for anonymity - author's prior published work] use the East and Central African country subset of this dataset.

countries, except for Madagascar due to its size, and did not include Somalia, because of the exceptionally limited fine-grained information available about armed group formation there. The coding protocol was designed to include even groups that failed after committing only minimal violence, if we could ascertain that the group planned to use violence to advance political goals. Due to the focus on initial group formation and violence, the coding protocol aimed to capture new anti-state armed groups; we excluded groups that had fragmented off of other, existing rebel groups or resulted from a merger of multiple, prior rebel groups.

By explicitly trying to capture all rebel groups that formed in a given place and time, and by scrutinizing the very initial phases of group formation and violence, this dataset captures new rebel group formation at a level of detail unmatched by prior datasets. About *one-third* of the 152 groups we capture are listed in commonly-used, prior datasets; only 31 percent of the rebel groups our research team identified are included in the UCDP/PRIO Non-State Actor list,<sup>22</sup> and only 32 percent are in the National Consortium for the Study of Terrorism and Responses to Terrorism's (START) Global Terrorism Database (GTD).<sup>23</sup> As I describe below, the dataset almost certainly undercounts the true number of rebel group that formed during this period.

---

<sup>22</sup> Other prominent datasets of rebel group attributes, such as Beber and Blattman (2013) and Braithwaite and Cunningham (2020), as well as the most widely-used dataset on ethnic rebellions (Cederman, Wimmer, and Min 2010), rely on the UCDP/PRIO Non-State Actor list to identify their population of rebellions.

<sup>23</sup> We included 11 groups in the dataset for which our certainty that they met our criteria was low. We coded groups as "low" certainty if we could identify only one reliable source indicating the existence of the group, or if there were many conflicting reports about the potential existence of a rebel groups, or if there was a political reason why a reported group may not have been a true group – especially if we identified political incentives for the government to falsely claim the existence of a rebel group. None of these low certainty groups were in UCDP/PRIO and only one can be found in GTD. However, even after dropping these group for my dataset, the percentage of rebel groups in the new dataset also found in UCDP rises only to 33% and GTD rises to 34%. All results in the paper include these 11 low certainty groups (although many get dropped using listwise deletion), and are robust to their exclusion.



Figure 1. Number of New Rebel Groups Formed (1997-2015)

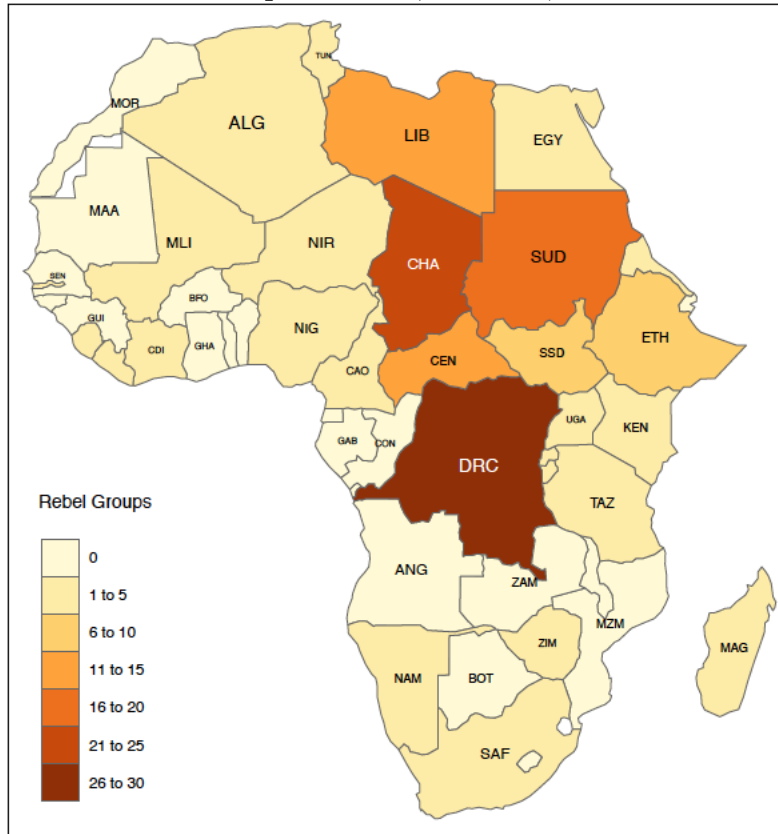


Figure 1 shows the state that each new rebel group targeted; 17% of rebel groups' initial operational planning activities (e.g. bases and training) were outside of the state they targeted. These quantities of rebel groups per state should be taken as a lower-bound. For all countries for which we identified more than 5 rebel groups forming during this period – as well as for Burundi and Madagascar – coders judged that there was a high likelihood of additional, unidentified new rebel groups that formed. Available information did not allow us to count additional groups with sufficient certainty for inclusion.

To identify potential rebel groups for inclusion, the research team started with actors listed in the Armed Conflict Location and Event Dataset (ACLED) (Raleigh et al. 2010) and consulted a wide variety of news media and publicly-available reports from non-governmental organizations and international organizations in English, French, and Swahili. We researched over one thousand potential groups. This resulted in a dataset of 152 armed groups. Of these 152 groups, 102 are listed in ACLED as a rebel group, an ethnic militia, or a political militia; the others are not listed in ACLED. We excluded from the dataset hundreds of actors that are named in ACLED because they did not meet our inclusion criteria. For example, many were

excluded because we determined that they formed prior to January 1997 or had fragmented from or were mergers of prior rebel groups. The Appendix discusses the inclusion criteria and research process in greater detail, and includes the full list of included rebel groups.

We also coded a series of binary variables on the organizational basis of each group (or lack thereof), groups' patterns of initial violence, initial resource endowments, whether they formed in a rural or urban area, and others. Coders were instructed to identify and compare among multiple sources to code variables wherever possible, and to leave variables missing where there simply was not enough information to make a reasoned judgement. Because information on these groups in U.S. library databases of news media and other publicly available on-line sources is often lacking in detail, especially about these early phases, there is considerable missing data on most of these variables. For example, the variable that aims to capture whether or not rebels were initially clandestine is missing for 36% of groups, and the variable that aims to capture whether groups initially benefited from natural resources (oil, diamonds, gold, or ivory) is missing for 20% of groups. A Coding Record document over 700 pages long that summarizes inclusion/exclusion decisions for all rebel groups considered, and coding decisions for included groups, including full citations, are posted on [author's website - a sample is provided in the Appendix for purposes of peer review].

There are some limitations of this data. While for most states, the research team deemed that we likely captured all or the vast majority of instances of rebel group formation, that was not possible for all countries and thus the dataset likely undercounts the true number. In one of the more extreme examples, the primary coder for Chadian rebel groups found that there were likely about 15 additional groups in Chad beyond the 22 that she included. The sources we could find lacked concrete information about these potential groups, preventing us from including them. In another example, the coder for Libya estimates that there were dozens if not

hundreds of new armed groups that formed there that may have met our criteria, but there is almost no information available about them, at least in English and French.<sup>24</sup> Echoing Weidman (2015)'s careful assessment of the accuracy of conflict event location data, another limitation we found when building the dataset -- is that it was often not possible to identify the precise *location* of rebel group formation without a great deal of speculation. Instead, to proxy for a subnational measure of state capacity, we coded approximate locations of a group's initial area of operational planning (e.g. location of training or bases) when possible, and whether this was in a rural or urban area. The lack of more fine-grained location information limits how the data can be used currently for subnational analysis. I return to this issue in the conclusion, arguing that using evidence from local sources to identify rebel groups' formation locations is an important next step for the frontier of conflict data collection.

## V. Analyses

### *Descriptive Statistics on Rebel Group Formation in Africa*

Here I report group-level, descriptive patterns in the data, which are summarized in Table 1 below. Consistent with the theoretical discussion above (*H1a*), the strong majority of African rebel groups that formed during this period did so in rural areas, not a major urban center (population roughly 500,000 or higher). Of the 152 rebel groups identified in the dataset for which we were able to code an initial location, 83% of them were in rural areas. This variable codes the physical location of operational planning for initial attacks in the rebels' target

---

<sup>24</sup> The ACLED information that we used relied on many Arabic sources. McQuinn (2015, 230), who estimates that there were 236 Libyan armed organizations operating in Misrata alone in 2011, based on his fieldwork there in 2011 and 2012.

country – the first identifiable operational planning location after the initial leadership decided to form a group that would commit violence against the state.<sup>25</sup> This is arguably a lower bound on the true proportion of rebel groups that formed in a rural area, where news media is scarce; there is missing data for 27% of rebel groups for this variable, as noted the “*n* of groups coded (out of 152)” column of Table 1.

The modal rebel group also initially (in their first year) appeared to be clandestine and did not benefit from substantial material endowments. Specifically, we determined that about 69% started as clandestine groups; media and other accounts indicate that they did not “advertise” their existence publicly for the first year after they formed an organization with plans to attack the state.<sup>26</sup> Only about 8% benefitted initially (in their first year) from revenue from oil, diamonds, ivory, or gold and only 25% benefitted initially (in their first year) from a foreign government that materially sponsored their activities via cash, weapons, training, or soldiers.

---

<sup>25</sup> This is not necessarily the same as location as where the initial attacks occurred although it was usually nearby. Any named location of operational planning, including mention of the group’s base or training activities, within a year of the decision to rebel sufficed.

<sup>26</sup> To ascertain whether a group was clandestine, we sought evidence that the group did not initially announce who they were and what they planned to do (e.g. not claiming responsibility for attacks that were later reliably attributed to the group), and that local civilians and the government did not initially attribute violence to the group. If that information was not available, we considered it evidence of clandestine intent if there was a lengthy (at least 3 month) period between when the group decided to become a violent group and when it conducted violence for which it claimed credit. If there was no evidence of a group concealing their existence/activities, nor any evidence of their operating out in the open, we coded the variable as missing.

Table 1. Initial attributes of rebel groups forming in Africa (1997–2015)

	% of rebel groups (estimate based on coded groups)	<i>n</i> of groups coded (out of 152)
Rural	83%	111
Clandestine	69%	97
Oil, diamonds, gold, ivory	8%	122
External sponsor	25%	128
Prior protest	11%	117
Prior contestation (protest or riots)	32%	116
Prior civil society org. (unions, student orgs, party, religious)	19%	133
Prior military organization	26%	122
Prior rebel organization	17%	112
No prior organization	33%	118
>25 battle-death attack in 1 <sup>st</sup> year	18%	128
>25 battle-death attack in 1 <sup>st</sup> 2 yrs	28%	130
>25 battle-death attack in 1 <sup>st</sup> 3 yrs	32%	130

Sources: Original dataset based on the Armed Conflict Location and Event Dataset (ACLED), news articles, and expert consultations. The variables capturing prior protest and prior contestation uses the Social Conflict in Africa Database (SCAD). The dataset includes 152 rebel groups. There was considerable missing data on some variables due to a lack of information in available sources; this is indicated in the “*n* of groups coded (out of 152)” column.

Turning to the mobilization and organizational antecedents of each rebel group, the research team sought to capture whether there was public, widespread contestation or other mobilization against the current government in the year prior to the formation of the organized armed (rebel) group the locality where the rebels formed. Protest mobilization can generate crucial organizational capacity for anti-state movements (Pearlman 2021). Our primary source for the contestation evidence was the Social Conflict in Africa Database (SCAD) (Salehyan et al. 2012). We coded two variables related to this concept: One capturing whether a protest occurred (an “organized demonstration” in SCAD), and another measuring contestation more inclusively (“organized demonstrations,” “disorganized demonstrations,” “organized



riots," or "disorganized riots" in SCAD). We operationalized "local" as either the same first-tier administrative unit (usually district or province) where the rebels formed, or a contiguous one. We operationalized "widespread" by counting events that appeared to involve a group of at least 100 people. By these standards, only 11% of rebel groups coded for this variable occurred where a major, local protest had occurred in the prior year (an "organized demonstration" in the SCAD dataset), and 32% formed in the year following a major, local contestation event, more broadly defined. There is moderate missing data (roughly 24%) for this variable, due largely due to the difficulties of identifying the localities where rebels appeared to form.

We also sought to capture whether the structure, processes, or membership of a prior formal organization was used *substantially* in building the rebel groups. We operationalized this variable as follows: Does the available evidence suggest that more than roughly one-third of rebel group membership previously (in the ten years prior to it taking up arms) was part of another organization? The data we were able to collect on these organizational endowments is not granular; we were not usually able to discern the degree or type of organizational reliance in much detail. Also, we usually had to rely on information about the membership of the rebel group in general; lack of detail in our sources prevented us from capturing this information specifically for the initial phases of rebel group formation. Using the data we were able to collect, we found that only about 19% of rebel groups relied on civil society organizations, such as political parties, religious, or other organizations such as unions or student organizations. A higher percentage of rebel groups appear to have relied on prior military organizations (26%), rebel groups (18%), or self-defense groups (19%). These categories are not mutually exclusive. We count 33% of groups as drawing no prior organizations, suggesting that this is a sizeable minority of groups.

Finally, only a fairly small minority of these rebel groups committed anything but small-scale violence at the outset. Only 18% committed an attack that resulted in at least 25 battle-related deaths in the first year after they formed. Only 28% did so in their first two years, and 32% did so in their first three years.<sup>27</sup> Internal armed conflict in Africa does not typically begin with an eruption of large-scale violence.

While the amount of missing data is high for some of these variables, missingness likely biases these numbers away from the account favored by this paper's arguments. Missing data is likely positively correlated with instances of rural group formation, due to relative media scarcity there. It is also likely negatively correlated with substantial resource and organizational endowments as well as committing early, large attacks, because of the conspicuous nature of these endowments and violence. A negative relationship between missingness and prior mobilization is also likely because missingness on the prior mobilization variables is driven by the lack of location information on early rebel location; groups that did form in the wake of major public contestation would more likely be captured in news media, given that the protests or riots would have already put the area in the public eye.

The direction of bias introduced by missingness on the organizational data is more contestable since rebels may hide their organizational ties for fear of revealing support networks, and doing so may be easier in a rural area. On the other hand, membership in prior nonviolent organizations is typically visible to the local population, and so is likely to be noted in reporting on a group; we included retrospective reporting on groups (not only sources from the outset of violence) to code this variable. For similar reasons, while we sought to only code ties to organizations where they were used substantially to build a rebel group – with more

---

<sup>27</sup> Note that this is not the same measure as the threshold for a conflict's inclusion in the UCDP/PRIO Armed Conflict Dataset; their inclusion criterion is 25 recorded battle-related total in a calendar year.

than about one-third of the rebels' membership coming from the prior organization – it is possible that this variable codes organizational ties in cases where they are, in fact, minimal. News reports may note ties to prior organizations when in fact just a small, subset subset of a rebel group's members belonged to a prior organization. This potential for under- or overcounting organizational ties indicates the importance of future efforts finding evidence that allows for scrutiny of to what extent, and how, armed groups build on prior organizations.

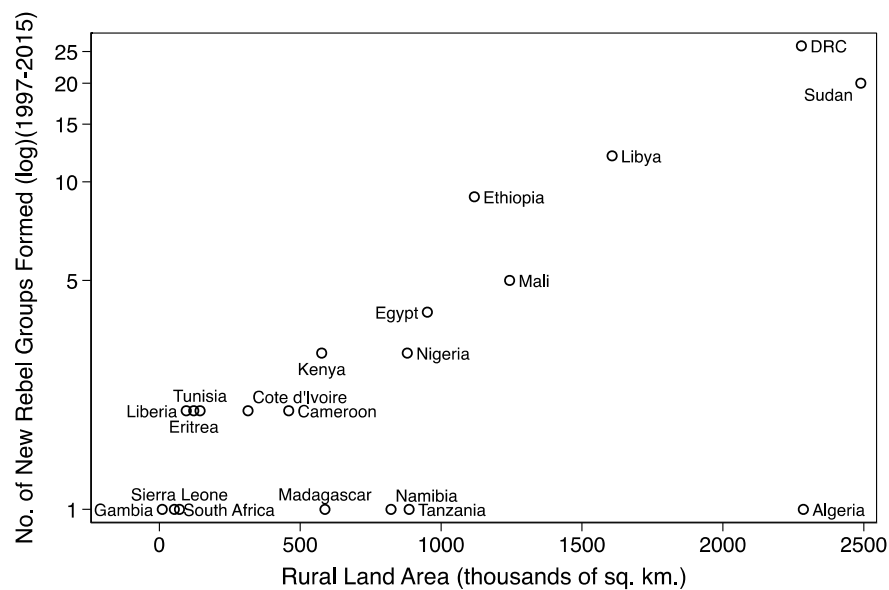
In sum, basic descriptive statistics from the new data support core propositions of the paper: The modal nascent rebel group in Africa since the late 1990s appears to have been initially rural, secretive, and poor; did not emerge in the wake of local public contestation; and committed only small-scale violence. The findings also reveal considerable variation in these initial conditions, underscoring the importance of examining the distinct drivers of different forms of anti-state armed conflict and its potential escalation. Keeping the limitations of the data in mind, the analyses below aim to explore the plausibility of the paper's propositions. Future works should use more fine-grained, complete data to dig deeper into these relationships.

#### *Exploratory Analyses on Rural Territory and Rebel Group Formation*

The paper has argued that states with more rural territory are more likely to experience more rebel group formation than those with less rural territory (*H1b*). As an initial cut at this relationship, I examine country-level data. A subnational analysis comparing rural versus urban territories' likelihood of becoming the location of a new rebel group would ideally supplement these cross-national analyses. However this is not presently possible with this data. As noted above, due to the challenges of obtaining reliable information about the precise location of rebels' early formation, and an effort to avoid coding variables based on thin sources and speculation, we captured only coarse measures of formation location.

To measure the extent of rural territory, I use a measure of rural land area from the Columbia University Center for International Earth Science Information Network (CIESIN) that uses satellite data and is based on a combination of population counts, settlement points, and the presence of nighttime lights.<sup>28</sup> A simple scatterplot of this measure against the number of rebel groups formed (logged) – Figure 2 below – clearly shows the expected, positive relationship.

Figure 2. Country-Level Rural Land Area and Rebel Group Formation



Cross-national regressions also show basic support for the first hypothesis (*H1a*).<sup>29</sup> I use static data given the relatively short period covered by the dataset and the lack of temporal

<sup>28</sup> Center for International Earth Science Information Network (CIESIN)/Columbia University. 2013. Urban-Rural Population and Land Area Estimates Version 2. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Accessed via World Bank Development Indicators Databank, August 2020.

<sup>29</sup> Lee and Zhang (2017) and Brambor et al (2019)'s measures of a state's informational capacity would be promising alternative measures for the concepts put forth in the theoretical argument; however, both focus explicit on the state's bureaucratic capacity to count their population, building on Scott (1998)'s theory of states making their populations legible. Those measures are not amenable for this analysis

variation in key variables. I use negative binomial models (models 3 and 4) because the dependent variable counts the number of rebel groups that formed in a country from 1997-2015, and is over-dispersed.<sup>30</sup> For ease of interpretation and to probe model dependence, I also show results using linear regression (models 1 and 2).

To examine whether this relationship may be spurious, I control for some of the most commonly proposed alternative explanations for rebellion. All measures come from the World Bank's World Development Indicators database. To control for the opportunity cost of rebelling, I use literacy rates among the population above age 15; results are similar using a measure of male unemployment to proxy for opportunity costs of rebellion. To control for the state's military capacity, I use a Stockholm International Peace Research Institute (SIPRI) measure of military expenditures as a percent of GDP (logged). To control for grievances induced by poor public service provision, I include a World Bank "Sustainable Energy for All" database measure that is based on household surveys of the percentage of the rural population with access to electricity (logged). I also include a standard control for population size (logged). I do not include a measure of country size since it is highly collinear with rural land area.

Models 1 and 3 show statistically significant, positive bivariate relationships between rural land area and rebel group formation. Models 2 and 4 show the model with controls; the coefficient on rural land area takes the expected (positive) direction and is statistically significant, lending some assurance that the bivariate correlations are not spurious. Using the negative binomial model in (8) and holding other covariates at their mean, increasing rural land area from the 50<sup>th</sup> percentile value (roughly that of Cote d'Ivoire or Congo-Brazzaville) to the

---

because the former is missing data for over half of African states and the latter includes only six African states.

<sup>30</sup> I do not use zero-inflated negative binomial models, despite a large portion of zeros in my DV, because there is little theoretical reason to believe that the zeros (countries with no rebel groups forming) are generated by a different process than the non-zeros. Results are robust to using Poisson models.

75<sup>th</sup> percentile value (roughly that of Egypt, Nigeria, or Tanzania) increases the expected number of rebel groups forming in a country by 1.16 rebel groups.

Table 2. Correlates of Rebel Group Formation in African States, 1997-2015

	(1)	(2)	(3)	(4)
	OLS	OLS	Neg. binom.	Neg. binom.
Rural land area (1,000s of sq. km.)	0.006*** (0.001)	0.006*** (0.001)	0.002*** (0.000)	0.002*** (0.000)
Literacy (% literate >age 15)		-0.002 (0.044)		-0.006 (0.012)
Military Expenditures (% of GDP)(log)		-0.765 (1.144)		0.059 (0.338)
Electrification (% of rural pop electricity)(log)		-0.186 (0.681)		0.139 (0.200)
Population size (log)		-0.089 (0.709)		0.190 (0.203)
Obs	35	34	35	34

Unit of analysis is the country. DV is the number of rebel groups that formed in a given country from 1997-2015.  
 \*\*\*p<0.01, \*\* p<0.05, \* p<0.10

These basic analyses support a relationship between African states' quantity of remote territory and their vulnerability to new rebel group formation, underscoring the difficulty of projecting authority over large distances. Future studies of rebel group formation should include the possibility that the extent of rural territory is in fact captures anti-state grievances under certain circumstances, or perhaps the state's intentional neglect or lack of capacity to respond to rebel group formation in remote areas.

### *Rebel Group-Level Evidence for H2*

I now turn to assessing evidence for *H2a* and *H2b*, which propose, respectively, that rebel groups that rely on prior mobilization like protests, or on pre-existing civil society organizations, will be more likely to form in urban areas than rural areas. The unit of analysis is the rebel group, and all covariates capture attributes of the rebel group summarized in Table 1. The main independent variables of interest – all described above and summarized in Table 1 – are the *protest*, *any contestation*, *no prior organization*, and *civil society* variables. Recall that all variables are dichotomous. Linear probability models are used to estimate all models shown for ease of interpretation; results are substantively and statistically similar using logit. Standard errors are clustered by country. Models 1, 3, 5, and 7 in Table 3 below examine bivariate relationships between each independent variable of interest and the dependent variable (rural or urban rebel group formation). Models 2, 4, 6, and 8 include several control variables as a check on whether the correlation is spurious. I use the clandestine variable as a control, since a desire to remain clandestine for a longer period could affect nascent rebels' choice of forming in a rural or urban context. Additionally, I control for the material endowments of the rebels – whether they benefitted substantially at the outset from resources from an external state, or from natural resources (oil, diamonds, gold, or ivory). These could also help groups overcome barriers to entry for rebellion, potentially confounding the relationship between mobilization processes and selection of an urban or rural site for rebellion.

Table 3. Correlates of Rural Rebel Group Formation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prior Protest	-0.267** (0.128)	-0.343** (0.132)						
Any Prior Contestation			-0.107 (0.085)	-0.180 (0.111)				
No Prior Organization					0.090 (0.087)	0.142 (0.102)		
Civil Society							-0.112 (0.110)	-0.114 (0.147)
Clandestine		0.069 (0.124)		0.051 (0.131)		0.060 (0.125)		0.038 (0.119)
External support		-0.116 (0.114)		-0.058 (0.110)		0.028 (0.080)		-0.021 (0.098)
Nat'l resources		-0.350* (0.202)		-0.318 (0.200)		-0.467** (0.197)		-0.460** (0.208)
<i>Obs</i>	99	59	99	59	58	58	99	61

DV is whether the rebels formed in a rural area. All use linear probability models for ease of interpretation; results are substantively similar, and more statistically significant, using logit. All models are estimated with standard errors clustered on country.

\*\*\*p<0.01, \*\* p<0.05, \* p<0.10

Models 1 and 2 indicate strong support for the expected, negative relationship between protest and rural rebel group formation (*H2a*). The relationship is statistically significant and substantively large, with rebel groups 34 percent less likely to form in a rural area if they emerge within a year of a local protest, based on Model 2. There is more moderate support for



the expected relationships with local contestation more broadly defined; the relationship is substantively large but not statistically significant at conventional levels ( $p$ -value = 0.12 for the *any contestation* variable in Model 4). Support for the organizational variables (*H2b*) is less strong: Models 6 through 8 show that the relationships are in the expected direction but are not substantively large or statistically significant. The control variables take the sign generally expected by the theoretical argument; groups that are initially clandestine are more likely to form in rural areas, and those with substantial material endowments are more likely to form in urban areas. This is consistent with the argument that a selection process governs rebel groups' decision whether to form in a rural or urban area, with groups that are better equipped to confront the state doing so in urban areas.

The number of observations in these models vary widely because missing data is present for all of these covariates, and all models in Table 3 take the standard approach of listwise deletion. As an alternative approach to handling missing data, I use multiple imputation in the appendix and re-estimate all models in Table 3 on the imputed data. The standard errors are somewhat larger in the resulting estimates, but the results are substantively similar.

## **VI. Conclusion**

New data introduced in this paper offers a more complete picture of new rebel group formation in Africa (1997-2015) than previous data allowed for, due to numerous missing instances of armed group formation in standard conflict datasets. Basic analyses show that new rebel group formation is more frequent than standard datasets capture, and that new rebel groups in Africa do not commonly form in large urban areas. The paper also finds some support for the proposition that nascent rebels in rural areas are less likely to emerge following

nearby nonviolent contestation – especially protests – than those that form in urban areas. I have argued that these findings are consistent with conceptualizing initial rebel group formation as an informational problem for states. States find it more challenging to monitor their rural territories for emerging security threats, and the absence of major public contestation leading up to rebel group formation may compound these detection challenges.

Much remains for future work. The paper’s descriptive findings suggest several open questions about the processes of rebel group formation in rural versus urban areas and implications for shifting patterns of political violence and mobilization as the Global South continues to urbanize; the early uses of violence and how it relates to mobilizing those outside an initial rebel vanguard; how and why only some small rebel groups go on to consolidate into larger groups; and how states pre-empt and respond to these dynamics. However, more detailed information on armed group formation – especially fine-grained information about rebel group formation location in Africa and beyond – will be needed for more rigorous analyses of these relationships and the causal mechanisms that drive them.

The data collection exercise underpinning this paper indicates the promise of examining armed group formation. But it also indicates the limits of what datasets can capture about the start of armed conflict in Africa “from a distance” – that is, without the benefits of local information gained from in-person fieldwork that benefitted prior, subnational studies in Africa (e.g. Weinstein 2007; Reno 2011; McQuinn 2015; Roessler 2016; Lewis 2020). These works and others on the microfoundations of conflict have demonstrated that fieldwork in the localities where rebels operate can recover basic facts about rebellion that are not possible to capture remotely. Partnerships with and support for local researchers, human rights organizations, and other local non-governmental organizations – all of which can more readily access local information about poorly-documented past conflict events – are another promising option, as

ACLED's local conflict observatories demonstrate.<sup>31</sup> Other initiatives that have successfully harnessed local sources of information to capture fine-grained data about conflicts include Sanchez de la Sierra (2020)'s technique of training local conflict reporters, and Humphreys and Van Der Windt (2015)'s "crowdseeding" approach to using technology to trace conflict-related events. A combination of these approaches carried out in several countries would be resource-intensive, but promises to greatly improve data quality on armed group formation, and to help bridge micro- and macro-level knowledge about the causes of civil conflict.

---

<sup>31</sup> See <https://acleddata.com/local-observatories/>.

## References

- Asal, Victor, Ken Cousins, and Kristian Skrede Gleditsch. 2015. "Making Ends Meet: Combining Organizational Data in Contentious Politics." *Journal of Peace Research* 52 (1): 134–38.
- Balcells, Laia. 2017. *Rivalry and Revenge: The Politics of Violence during Civil War*. Reprint edition. S.L.: Cambridge University Press.
- Balcells, Laia, and Patricia Justino. 2014. "Bridging Micro and Macro Approaches on Civil Wars and Political Violence: Issues, Challenges, and the Way Forward." *Journal of Conflict Resolution* 58 (8): 1343–59. <https://doi.org/10.1177/0022002714547905>.
- Bates, Robert H. 1983. *Essays on the Political Economy of Rural Africa*. Cambridge: Cambridge University Press.
- Blattman, Chris, and Bernd Beber. 2013. "The Logic of Child Soldiering and Coercion." *International Organization* 67 (1): 65–104.
- Blaxland, Joel. 2020. *Insurgency Prewar Preparation and Intrastate Conflict*. Springer International Publishing. <https://link.springer.com/book/10.1007/978-3-030-38185-1>.
- Boone, Catherine. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice*. Cambridge, UK: Cambridge University Press.
- Braithwaite, Jessica Maves, and Kathleen Cunningham. 2020. "When Organizations Rebel: Introducing the Foundations of Rebel Group Emergence (FORGE) Dataset." *International Studies Quarterly*.
- — —. Forthcoming. "When Organizations Rebel: Introducing the Foundations of Rebel Group Emergence (FORGE) Dataset." *International Studies Quarterly*.
- Braithwaite, Jessica Maves, and Amanda Abigail Licht. 2020. "The Effect of Civil Society Organizations and Democratization Aid on Civil War Onset." *Journal of Conflict Resolution* 64 (6): 1095–1120. <https://doi.org/10.1177/0022002719888684>.
- Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2019. "The Lay of the Land: Information Capacity and the Modern State." *Comparative Political Studies*, May. <https://doi.org/10.1177/0010414019843432>.
- Branch, Adam, and Zachariah Cherian Mampilly. 2015. *Africa Uprising: Political Protest and Political Change*. African Arguments. London: Zed Books.
- Brass, Jennifer N. 2016. *Allies or Adversaries: NGOs and the State in Africa*. New York NY: Cambridge University Press.
- Butcher, Charles, Jessica Maves Braithwaite, Jonathan Pinckney, Eirin Haugseth, Ingrid Vik Bakken, and Marius Swane Wishman. 2021. "Introducing the Anatomy of Resistance Campaigns (ARC) Dataset." *Journal of Peace Research*. <https://www.arc-project.net/publications>.
- Butcher, Charles, and Isak Svensson. 2016. "Manufacturing Dissent: Modernization and the Onset of Major Nonviolent Resistance Campaigns." *Journal of Conflict Resolution* 60 (2): 311–39.
- Calle, Luis de la, and Ignacio Sánchez-Cuenca. 2012. "Rebels without a Territory: An Analysis of Nonterritorial Conflicts in the World, 1970–1997." *Journal of Conflict Resolution* 56 (4): 580–603.
- Cederman, Andreas Wimmer, and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62 (1).
- Chenoweth, Erica, and Maria J. Stephan. 2013. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Paperback ed. Columbia Studies in Terrorism and Irregular Warfare. New York, NY: Columbia Univ. Press.

- Coleman, James S. 1977. "The Concept of Political Penetration." In *Government and Rural Development in East Africa: Essays on Political Penetration*, edited by L. Cliffe, J. S. Coleman, and M. R. Doornbos. The Hague: Martinus Nijhoff.
- Cunningham, David E., Kristian Skrede Gleditsch, Belén González, Dragana Vidović, and Peter B. White. 2017. "Words and Deeds: From Incompatibilities to Outcomes in Anti-Government Disputes." *Journal of Peace Research* 54 (4).
- Cunningham, David, and Lemke. 2014. "Beyond Civil War: A Quantitative Examination of Causes of Violence within Countries." *Civil Wars* 16 (3): 328–45.
- Cunningham, Kathleen. 2013. "Understanding Strategic Choice The Determinants of Civil War and Nonviolent Campaign in Self-Determination Disputes." *Journal of Peace Research* 50 (3): 291–304.
- Dawkins, Sophia. 2020. "The Problem of the Missing Dead." *Journal of Peace Research*, December, 0022343320962159. <https://doi.org/10.1177/0022343320962159>.
- Dietrich, Nick, and Kristine Eck. 2020. "Known Unknowns: Media Bias in the Reporting of Political Violence." *International Interactions* 0 (0): 1–18. <https://doi.org/10.1080/03050629.2020.1814758>.
- Elkhamri, Mounir, Lester W. Grau, Laurie King-Irani, Amanda S. Mitchell, and Lenny Tasa Bennett. 2005. "Urban Population Control in a Counterinsurgency." Fort Leavenworth, Kansas: United States Army Foreign Military Studies Office. <https://apps.dtic.mil/sti/citations/ADA434983>.
- Fearon, James, and David Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review* 97 (1): 75–90.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39 (5): 615–37.
- Golooba-Mutebi, Frederick, and Anders Sjögren. 2017. "From Rural Rebellions to Urban Riots: Political Competition and Changing Patterns of Violent Political Revolt in Uganda." *Commonwealth & Comparative Politics* 55 (1): 22–40. <https://doi.org/10.1080/14662043.2017.1248215>.
- Gyimah-Boadi, Emmanuel. 1997. "Civil Society in Africa." In *Consolidating the Third Wave Democracies*, edited by Larry Diamond, Marc F. Plattner, Yun-han Chu, and Hung-mao Tien. Baltimore, Md: Johns Hopkins University Press.
- Haim, Dotan. 2019. "Civilian Social Networks and Credible Counterinsurgency." *Working Paper*.
- Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, N.J.: Princeton University Press.
- Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge and New York: Cambridge University Press.
- Kasfir, Nelson. 1998. "Civil Society, the State and Democracy in Africa." *Commonwealth & Comparative Politics* 36 (2): 123–49. <https://doi.org/10.1080/14662049808447770>.
- — —. 2005. "Guerillas and Civilian Participation: The National Resistance Army in Uganda, 1981-1986." *Journal of Modern African Studies* 43 (2): 271–96.
- Kilcullen, David. 2010. *Counterinsurgency*. Oxford ; New York: Oxford University Press.
- Kocher, Matthew. 2004. *The Human Ecology of Civil War*. University of Chicago: Ph.D. Dissertation.
- Koren, Ore, and Anoop K Sarbahi. 2018. "State Capacity, Insurgency, and Civil War: A Disaggregated Analysis." *International Studies Quarterly* 62 (2): 274–88. <https://doi.org/10.1093/isq/sqx076>.

- Lee, Melissa, and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *Journal of Politics* 79 (1): 118–32.
- Leites, Nathan, and Charles Wolf. 1970. *Rebellion and Authority: An Analytic Essay on Insurgent Conflicts*. Markham Series in Public Policy Analysis. Chicago: Markham Pub. Co.
- Leventoğlu, Bahar, and Nils W. Metternich. 2018. "Born Weak, Growing Strong: Anti-Government Protests as a Signal of Rebel Strength in the Context of Civil Wars." *American Journal of Political Science* 62 (3): 581–96.
- Lewis, Janet I. 2017. "How Does Ethnic Rebellion Start?" *Comparative Political Studies* 50 (10): 1420–50.
- — —. 2020. *How Insurgency Begins: Rebel Group Formation in Uganda and Beyond*. Cambridge Studies in Comparative Politics. Cambridge University Press.
- Malone, Iris. 2019. *Insurgency Formation and Civil War Onset*. Stanford University: Ph. D. Dissertation.
- — —. Forthcoming. "Unmasking Militants: Organizational Trends in Armed Groups, 1970–2012." *International Studies Quarterly*.
- Mazur, Kevin. 2021. *Revolution in Syria: Identity, Networks, and Repression*. Cambridge Studies in Comparative Politics. Cambridge, United Kingdom ; New York, NY: Cambridge University Press.
- McQuinn, Brian. 2015. "History's Warriors: The Emergence of Revolutionary Brigades in Misrata." In *The Libyan Revolution and Its Aftermath*, edited by Peter Cole and Brian McQuinn. London: Hurst Press.
- Mkandawire, Thandika. 2002. "The Terrible Toll of Post-Colonial 'Rebel Movements' in Africa: Towards an Explanation of the Violence against the Peasantry." *The Journal of Modern African Studies* 40 (2): 181–215.
- Moore, Barrington. 1984. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Repr. London: Allen Lane.
- Mueller, Lisa. 2018. *Political Protest in Contemporary Africa*. Cambridge University Press.
- Museveni, Yoweri. 1997. *Sowing the Mustard Seed: The Struggle for Freedom and Democracy in Uganda*. London: Macmillan.
- O'Donnell, Guillermo. 1993. "On the State, Democratization, and Some Conceptual Problems: A Latin American View with Some Glances at Post-Communist Countries." *World Development* 21 (8): 1355–69.
- Parkinson, Sarah E., and Sherry Zaks. 2018. "Militant and Rebel Organization(s)." *Comparative Politics* 50 (2): 271–93.
- Parkinson, Sarah Elizabeth. 2013. "Organizing Rebellion: Rethinking High-Risk Mobilization and Social Networks in War." *American Political Science Review* 107 (3): 418–32. <https://doi.org/10.1017/S0003055413000208>.
- Pearlman, Wendy. 2021. "Mobilizing From Scratch: Large-Scale Collective Action Without Preexisting Organization in the Syrian Uprising." *Comparative Political Studies* 54 (10): 1786–1817. <https://doi.org/10.1177/0010414020912281>.
- Raleigh, Clionadh, Andrew Linke, Havard Hegre, and Joakim Karlsen. 2010. "Introducing ACLED-Armed Location and Event Data." *Journal of Peace Research* 47 (5): 1–10.
- Reno, William. 2011. *Warfare in Independent Africa*. Cambridge, UK ; New York: Cambridge University Press.
- Roessler, Philip G. 2016. *Ethnic Politics and State Power in Africa: The Logic of the Coup-Civil War Trap*. Cambridge, United Kingdom ; New York, NY: Cambridge University Press.

- Ryckman, Kirssa Cline. 2020. "A Turn to Violence: The Escalation of Nonviolent Movements." *Journal of Conflict Resolution* 64 (2-3): 318-43. <https://doi.org/10.1177/0022002719861707>.
- Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38 (4): 503-11. <https://doi.org/10.1080/03050629.2012.697426>.
- Sanchez de la Sierra, Raul. 2020. "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo." *Journal of Political Economy*.
- Scott, James C. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven (Conn.): Yale University press.
- — —. 1998. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Nachdr. Yale Agrarian Studies. New Haven, Conn.: Yale Univ. Press.
- Sinno, Abdulkader H. 2008. *Organizations at War in Afghanistan and Beyond*. Ithaca: Cornell University Press.
- Staniland, Paul. 2010. "Cities on Fire: Social Mobilization, State Policy, and Urban Insurgency." *Comparative Political Studies* 43 (12): 1623-49. <https://doi.org/10.1177/0010414010374022>.
- — —. 2014. *Networks of Rebellion: Explaining Insurgent Cohesion and Collapse*. Cornell Studies in Security Affairs. Ithaca: Cornell University Press.
- Straus, Scott. 2012. "Wars Do End! Changing Patterns of Political Violence in Sub-Saharan Africa." *African Affairs* 111 (443): 179-201. <https://doi.org/10.1093/afraf/ads015>.
- Thurber, Ches. 2021. *Between Mao and Gandhi: The Social Roots of Civil Resistance*. Cambridge University press.
- Tilly, Charles. 1978. *From Mobilization to Revolution*. Reading, Mass: Addison-Wesley Pub. Co.
- Van der Windt, Peter, and Macartan Humphreys. 2015. "Crowdseeding Conflict Data." *Journal of Conflict Resolution* Forthcoming.
- Weidmann, Nils B. 2015. "On the Accuracy of Media-Based Conflict Event Data." *Journal of Conflict Resolution* 59 (6): 1129-49. <https://doi.org/10.1177/0022002714530431>.
- Weinstein, Jeremy M. 2007. *Inside Rebellion*. Cambridge: Cambridge University Press.
- Zukerman-Daly, Sarah. 2012. "Organizational Legacies of Violence: Conditions Favoring Insurgency Onset in Colombia, 1964-1984." *Journal of Peace Research* 49 (3): 473-91.