4 Research Design

Developing Empirical Tests of the Theory

> Don't compare me to the Almighty. Compare me to the alternative.
>
> Joseph R. Biden, President of the United States

> A world without the United Nations... would be more costly to all and more dangerous.
>
> Richard Holbrooke, US Ambassador to the UN, 1999–2001

Using a formal model, Chapter 3 introduced localized peace enforcement theory. The framework of the theory examines the conditions under which communal violence breaks out in areas where peacekeeping forces are deployed. Recognizing local populations' preexisting beliefs about the biases of international actors, I shed light on the influential role that domestic perceptions of peacekeepers play in evaluating the ability of peacekeepers to maintain peace at the community level. The enduring legacies of colonialism strongly shape these perceptions. My central argument posits that peacekeepers, when domestic populations believe they are impartial, can change individual beliefs about the likelihood that others will reciprocate their attempts to cooperate – thereby increasing individuals' willingness to cooperate and decreasing the likelihood that communal violence will break out.

Localized peace enforcement theory suggests three mechanisms through which impartial peacekeepers reduce communal violence. Each is formulated as a set of hypotheses (see Figure 3.1). This chapter outlines the research design employed to test those hypotheses. Generally speaking, the book's design is informed by the idea informally expressed by President Joe Biden in the chapter's first epigraph: Policies should be evaluated against the most likely (rather than ideal) other option. I therefore investigate whether a "world without the United Nations" would indeed be more dangerous, as Richard Holbrooke suggests.

I first investigate the deployment of peacekeepers of different nationalities to address communal violence within a single country, Mali, from 2013 to 2020. I selected Mali as a case study because it generalizes well to other settings with widespread communal violence and a multidimensional UN peacekeeping operation (PKO). Furthermore, the simultaneous UN and French deployments allow me to compare the

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

effectiveness of the UN to that of a former colonial power that is likely to be perceived as biased. With different types of personnel from member states all over the world, there is also significant variation within UN peacekeeping in Mali, which allows me to examine the effectiveness of different contributors. Finally, Mali is a least-likely test of my theory for at least three reasons. First, it is one of sub-Saharan Africa's poorest and most underdeveloped states, and its government lacks the institutional capacity to enforce violations of intergroup cooperation. Second, the perpetual salience of communal violence along multiple group dimensions makes intergroup cooperation a difficult proposition. Third, the convergence of French and UN peace operations should make local-level UN peacekeeping less effective in Mali than in other settings according to my theory because locals may confuse UN peacekeepers and French soldiers.

The complex nature of local-level PKOs necessitated a pluralistic and nimble data collection strategy in Mali. On the one hand, I required rich and in-depth information from conflict and postconflict settings with PKOs to understand how individuals think of the UN peacekeepers in their midst. On the other hand, I wanted to see how the deployment of peacekeepers shaped actual levels of violence across time and space. In addition, PKOs by nature select into violent settings, complicating any naive comparison of settings with and without peacekeepers since locations with peacekeepers would also have more conflict. At each level of analysis, I needed to account for confounding variables in order to make causal inference possible.

I therefore combined multiple methods of data collection at different levels of intensity and geographic scopes. My goal was for the structure of the empirical analysis to mirror the book's primary argument – to show how peacekeeping works from the bottom up, from the individual to the community to the country as a whole. Using individual- and subnational/community-level data from Mali as well as cross-national data from the universe of multidimensional PKOs deployed in Africa, I employ a three-part strategy to assess these hypotheses in the next few chapters of the book. First, I test the micro-level behavioral implications of the theory (H1a–H1b and H2a–H2c) using a lab-in-the-field experiment and a survey experiment, both implemented in Mali. Second, I test whether UN peacekeepers' ability to increase individual willingness to cooperate aggregates upward to prevent communal violence in Mali (H3). Third, I consider whether these findings extend to other countries. This chapter describes the data collection strategy and research method employed for each type of empirical analysis.

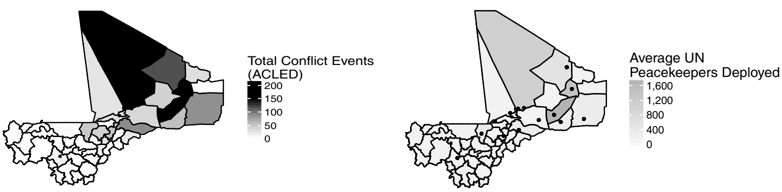
The experimental methods I use to analyze peacekeeping in Mali form part of a broader multimethod approach that experimentally tests

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

the observable implications of hypotheses deduced from theoretical first principles. This approach aligns with what Susan Hyde has called “theory first, experiments second” (2015): “the value of an experiment relative to a particular research agenda in [International Relations] can become clear when lab, survey, and/or field experiments are integrated into a broader project in which the observable implications of a theory are clearly spelled out; when experiments are clearly connected to these observable implications; and when experimental findings are replicated across theoretically relevant contexts” (p. 412). While the book relies on these experiments to make causal inferences about specific outcomes, the experiments of course do not exist in isolation from outcomes related to war and peace. I take seriously Aila Matanock’s admonitions about experimental work on conflict and postconflict settings. She maintains that new studies in peacebuilding should more explicitly link to theoretical frameworks, including by incorporating more outcome variables that are explicitly related to conflict recurrence, and feature coordinated sets of experiments using common treatments linked to arguments about how peace consolidates across contexts (2020, p. 2).

Lab-in-the-Field Experiment

Testing the argument that peacekeepers increase individuals’ willingness to cooperate with members of out-groups requires isolating the effect of peacekeeping from two primary threats to identification. The first threat comes from characteristics of the local context that may inhibit or bolster cooperation. For example, peacekeepers deploy to violent and unstable areas where intergroup cooperation is difficult to sustain, which generates a spurious negative correlation between peacekeeping



(a) Total conflict events in Mali (b) Average UN deployment in Mali

Figure 4.1 Peacekeepers deploy to the most violent settings
Note: Conflict events from the Armed Conflict and Event Location Dataset (ACLED) (Raleigh et al. 2010). Peacekeeping deployment numbers from the Robust Africa Deployments of Peacekeeping Operations (RADPKO) (Hunnicutt and Nomikos 2020); darker areas indicate more peacekeepers and black dots denote UN bases.

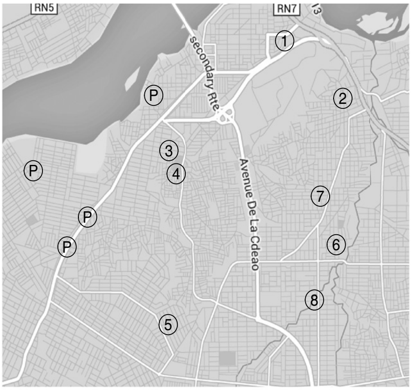
https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

and cooperation. As Figure 4.1 illustrates, this issue is particularly severe in Mali, as the UN deploys to the most violent areas of the country. Figure 4.1(a) visualizes the total number of conflict events by administrative district in Mali. Figure 4.1(b) displays the average number of UN peacekeepers deployed in those same areas.

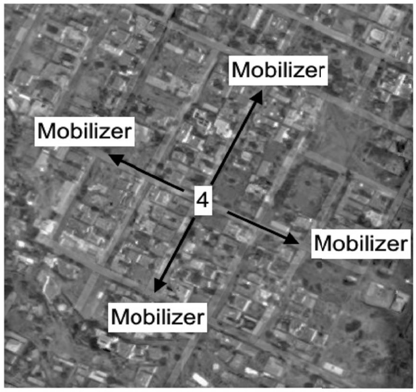
Second, the effect of peacekeeping patrols enforcing an interaction must be isolated from the impact of information generated by the circumstances surrounding that interaction. Each exchange between members of different social groups introduces new information that likely affects whether an individual will want to cooperate or not, which makes it difficult to disentangle the effects of this new information from those of peacekeeping enforcement. For instance, cattle herders may meet members of their community in a weekly market patrolled by peacekeepers but only choose to sell the meat from some of their cows after several peaceful interactions. Using solely observational data, we could come to the premature conclusion that the peacekeeping patrols increased the cattle herders' willingness to cooperate. Yet, it would not be clear whether the herders cooperated because of the presence of the peacekeepers or due to new information about their fellow community members gathered during the weekly market.

To address these two threats to identification, I embedded a randomized experiment within a "trust game" conducted in Mali. As part of the game, I tasked non-Tuareg Malians with sending part of their initial salary to a Tuareg Malian partner, who they then had to trust would reciprocate the attempt at cooperation by returning some of that contribution.[1] To investigate the first threat to identification – that local factors may inhibit or bolster cooperation – I randomly assigned participants to a control group or one of two peacekeeping treatments so that any potential characteristics of the area of the study (i.e., peri-urban neighborhoods of Mali) would be independent of the effect of enforcement. I informed participants in the treatment groups that a peacekeeper, either from the UN or France, would observe and impose fines on low contributions, a common method in experimental psychology to operationalize the presence of a third-party enforcer (Fehr and Fischbacher 2004; Bernhard, Fehr and Fischbacher 2006). To deal with the second threat to identification – that peacekeeping patrols enforce cooperation *and* generate new information during interactions with civilians – I kept the characteristics of the interaction between the participants the same

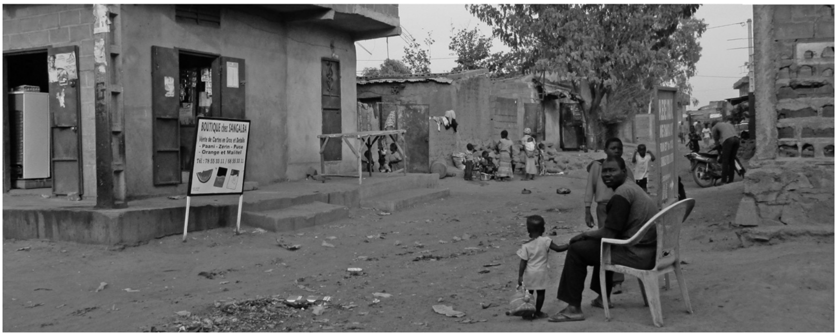
[1] Berg, Dickhaut and McCabe (1995) are typically credited with devising the first such trust game. They sought to explain why individuals would send anything at all in these games, which contradicts traditional rational choice assumptions in game theory. Social scientists have since sought to explain the outcomes of these games, which Elinor Ostrom called "better than rational" for participants (see Ostrom and Walker (2003)).



(a) Randomly selected neighborhoods



(b) Sampling procedure in Neighborhood 4



(c) Neighborhood 7. Photograph by author.

Figure 4.2 Sampling procedure for the lab experiment
Attribution notes: Imagery ©2024 Airbus, CNES/Airbus, Maxar Technologies, Map data ©2024 Google.

across treatment groups to ensure that any differences between the treatment and control groups can only be attributed to the effect of the treatment groups, rather than to new information that arises during the social encounter.[2]

Sample

I drew a sample of 512 non-Tuareg Malians from eight randomly selected neighborhoods of southeast Bamako, a semirural residential area of the capital city. As visualized in Figure 4.2(a), I randomly selected

[2] Ana Bracic argues that lab experiments are especially effective at measuring behavior toward out-groups in settings where norms against discrimination may lead to preference falsification (2016).

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

twelve GPS coordinates representing neighborhoods: Four for piloting (indicated with a P) and eight for the experiment (numbered). I chose this part of Bamako to minimize the differences between the capital and other areas of the country, which can mostly be found in the urban areas of Bamako not sampled here. The mobilizers brought the participants to a central location where one of eight enumerators met them and explained the rules of the game. A local field manager gave each participant a detailed briefing before the game and a debriefing afterward to ensure full comprehension.

I chose Bamako as the setting for the lab experiment for three reasons. First, it is the center of international PKOs in Mali: The United Nations Multidimensional Integrated Stabilization Mission in Mali (MINUSMA) has its main peacekeeping headquarters and central police operations there. Participants are therefore likely to be familiar with the UN peacekeeping mission; responses to the survey confirmed that this is the case. More than two-thirds (68 percent) reported seeing UN peacekeepers "all the time" or "often." Only 2 percent reported never having seen them. Bamako therefore offers the lab experiment a high degree of internal validity: Given respondents' awareness of the UN, the observed treatment effects are likely to operate as theorized. The second reason I conducted the lab experiment in Bamako is that since 2012, violence in northern and central Mali has forced internally displaced Tuareg to take up residence there, which has further diversified the ethnic makeup of the city's neighborhoods and increased the incidence of communal disputes. The third reason is that the November 2015 attack on the Radisson Blu hotel in Bamako – in which extremists killed twenty people, the first of its kind since the June 2015 peace accords – increased the salience of violence for respondents at the time of the experiment (BBC News). Shortly after the experiment ended on March 21, 2016, there was another armed group attack on the EU training mission in Bamako. Moreover, the frequent attacks on UN peacekeepers and the surrounding population generate ubiquitous headlines in Bamako. Together, these factors make the city an appropriate test case of local-level peacekeeping in Mali.

This sample is representative of residents in Bamako. However, it is not representative of all residents of Mali in two respects. First, even in these residential areas, the proximity to the center makes life substantially different than it is for the rural population. Second, the ethnic makeup of the sample differs between Bamako and the rest of Mali, specifically the percentage of the population that is Tuareg. However, because the observable implications concern only the behavior of non-Tuareg Malians toward Tuareg Malians, this does not affect the outcomes of interest.

The average participant was twenty-six years old, has two children, and has completed middle school. Only half of the participants were Bamako natives, and roughly half have a close friend who is Tuareg. About 34 percent of the participants reported belonging to the Bambara group, the country's largest ethnic group, which is similar to levels across Mali according to the most recent census. The sample was thus largely representative of the country's population with the exception of the gender balance. It was very difficult to recruit women for the experiment; due to cultural considerations, very few were willing to leave their neighborhood to participate in the experiment (described in more detail in the next section). As a remedy, I sent enumerators directly to respondents' homes for the survey experiment presented in Chapter 6. I fielded that survey over the course of two rounds in August and December 2017, giving me a chance to incorporate lessons learned from the sampling here. The results of the survey experiment suggest that the findings based on the predominantly male sample in the lab-in-the-field experiment generalize to a more gender-balanced sample. Moreover, given that men are overwhelmingly the perpetrators of communal violence in West Africa, determining the effect of peacekeeping on a predominantly male sample is critical for understanding the determinants of peace. Underscoring the gendered nature of communal violence, Krause writes in a study of a nonviolent community in Nigeria that "men from there were often mocked for not having fought [by men who said] 'they are women'" (2019*a*, p. 1466).

Figure 4.2(c) illustrates that the GPS coordinates I selected are not heavily trafficked areas or urban centers. For each neighborhood, I traveled to the coordinates and found the closest four-point intersection. I took pictures of this intersection so that my local mobilizers could easily find it. They worked in groups of four, moving in all four directions of the intersection until they came across a major road or obstacle, at which point they took a right turn (see Figure 4.2(b) for an example). This process continued until each mobilizer recruited their daily quota. Although the mobilizers did not have a specific recruiting protocol beyond these instructions, they were told to gather subjects from diverse backgrounds.

Research Safety and Ethical Considerations

The safety of the participants and local enumerators was of utmost importance to me.[3] I took every available precaution to avoid exposing my subjects to any unnecessary risk, judgment, or punishment from a peacekeeper based on their behavior in my study. Since the empirical

[3] Ethics are especially important in conflict settings. See Campbell (2017).

strategy of the game, which focuses on the one-shot contributions of non-Tuareg Malians, did not involve any response from the peacekeeper, I did not invite real peacekeepers to participate in the experiment. Thus, the participants did not observe any peacekeeper behavior. They were merely told that two peacekeepers from either France or the UN, depending on the treatment group, would fine them if they made a low contribution. The game ended before they would expect to see the result of the (imaginary) fine. Since peacekeepers lack the capacity to punish every violation, they rely on deterrence. Having a game with anticipated rather than actual punishment thus reflects how peacekeepers use an important coercive instrument.

I did not use real Tuareg partners because the volatile nature of intergroup relations in Mali, an active conflict setting, could have produced hostile feelings toward Tuareg players. It would have been difficult to ensure the safety of all participants if such hostilities escalated. Moreover, the hypotheses generate testable implications about the willingness of members of one group to cooperate with others. Because data collection did not require the involvement of the target of cooperation, the participants were not exposed to any risk. Finally, I made sure that the preprogrammed return from the Tuareg at the end of the game was generous to perpetuate positive impressions of Tuareg Malians. All subjects received the same monetary payoff upon exiting the study (1,500 FCFA).

All participants were thoroughly debriefed after the experiment until it was clear that they understood the nature of the game. No one indicated any concerns about the experiment, and most expressed satisfaction at being part of a study that contributed to what one participant characterized as "peace and reconciliation in Mali." A related issue is that participating in a lab experiment that involves deception might change participants' behavior in future studies. Since this is a worry for researchers rather than the individuals involved, the key concern is whether the deception was worth the cost of potentially changing the behavior of participants in future research. In his comparison of the use of deception in psychology (which frequently uses deception) and economics (which rarely does), Dickson (2011) explains that the salient alternative is "an otherwise identical experimental design in which the same stimuli are presented to subjects, but explicitly labeled as 'hypothetical' ... In judging the potential usefulness of deception, then, a natural question to ask is whether an individual's mode of psychological engagement with a stimulus depends on whether that stimulus is framed as being 'real' as opposed to hypothetical. If the answer to this question is 'yes' – and if this would make a substantial enough difference for measurements of the quantities of interest – then at the least a benefit from deception will have been identified" (p. 120). Given that the use of deception offers me, as a researcher, the opportunity to *safely* analyze the potential factors leading to intergroup cooperation in a setting torn apart by a lack of cooperation, I firmly believe there is a sufficient research-related benefit to cooperation as well.

Survey Experiment

To further test my theory's individual-level implications, I conducted a survey experiment, a format that is useful for at least four reasons. First, it provides further micro-level evidence of the proposed mechanism that perceptions of bias shape the likelihood that a peacekeeper will succeed. Second, it overcomes some key shortcomings of the lab-in-the-field experiment. The lab experiment is well suited to identifying and testing the mechanism under controlled circumstances, but raises the potential concern that the findings are not generalizable outside the lab. The survey experiment suggests that they *are* generalizable: The results indicate that impartial peacekeepers increase local residents' willingness to cooperate with a member of a different social group and positively affect beliefs about the prospects of peacefully resolving a communal dispute. Third, the survey experiment allows me to vary the conditions under which a dispute breaks out, which permits a direct comparison of the various mechanisms hypothesized to explain what makes a dispute likely to turn violent. Finally, the survey allows me to measure broad outcomes related to perceptions rather than behavior alone.

Ethical Considerations

To ensure the safety of the participants and local enumerators, I developed the survey, recruitment protocols, and implementation procedure in consultation with local partners from the areas sampled in the experiment (Davis 2020). I took care throughout the research process to respect the autonomy of the survey participants. I made it clear that the survey was part of a research study and allowed participants to ask the enumerators questions about the study at any time. The greatest ethical issue was asking questions that may be triggering for individuals. I consulted with local collaborators, as Justine Davis (2020) recommends in her discussion of best practices for conducting research in Africa. Where there was any doubt, I simply removed the question, as was the case with several questions about exposure to violence.

I have a lot of experience conducting research in Mali, and I understand that there are power differentials between myself and the respondents. Due to concerns that this could make them feel pressured to participate or affect their evaluations of the risks and benefits of doing so, each survey was administered by an enumerator from the same region, either Bamako or Segou, as the respondents. The enumerators were overseen by a Malian field manager who reported directly to me.

The enumerators explained the purpose of the study to each participant using a standardized consent document that I wrote in French, which is largely understood by the overwhelming majority of residents. To secure fully informed consent, the enumerators could clarify parts of the consent process or the survey in local languages. The respondents understood that they were participating in a research study and were allowed to opt out at any point. No one chose to do so, and the vast majority of respondents shared their opinions enthusiastically. Enumerators debriefed respondents after the survey, and provided them with local contact information if they had any follow-up questions. None of the respondents followed up.

All participants were assured of confidentiality during the consent process. Enumerators collected first names and phone numbers (where available) in case follow-up was necessary (it was not). They encrypted this information and sent it to me, and deleted it from their tablets at the end of the survey period. I separated any personal identifying information from the responses and saved each set of data to a spreadsheet. Only I have access to the spreadsheet with the personal information. To the best of my knowledge – and that of the Institutional Review Board and local collaborators who reviewed the research – the study caused no harm or trauma, including to research assistants and staff.

Sampling Procedure

In the first round (July–August 2017), I randomly selected eight peripheral neighborhoods in the capital of Bamako using the same four-corner procedure described for the lab experiment sampling (Figure 4.3a). Bamako is an important setting for local-level peacebuilding for the reasons described in Chapter 5. In the second round of the survey (December 2017),[4] I sampled 360 respondents from twelve villages in the Markala and San *communes* (districts) of the central Malian province of Segou (Figure 4.3b).[5] I chose six villages from Markala *commune*, which France invaded during the 2013 military intervention, and six from San *commune*, which it did not invade. I selected villages that had

[4] Outbreaks of communal violence in Segou forced me to postpone the second round to ensure the safety of the research team and participants.

[5] Due to the existence of high levels of local-level violence in Segou, the Malian government has outlawed travel by motorbike in the region. This law, combined with the travel distances between villages and the lack of navigable roads in Segou, made data collection much more challenging there than in Bamako.

(a) Neighborhoods in Bamako (Round 1) (b) Villages in Segou Region (Round 2)

Figure 4.3 Sampling for the survey experiment
Note: Fifteen neighborhoods were randomly selected in Bamako (numbers 1–4 were used for piloting; 5–12 were used for the survey experiment).

more than 500 residents in the latest census (2009) and obtained a set of six pairs of villages matched according to population size and the number of schools, hospitals, and clean water wells. Four enumerators under my supervision conducted tablet-assisted, in-person interviews.

The second-round sample improved the quality and generalizability of the survey findings in three ways. First, the sample comprised rural respondents, who serve as an important contrast to the mostly urban residents of Bamako. Second, Segou was a recent postconflict area in which it was safe to conduct research. While Islamic extremists seized control of the province during the heavy fighting of 2012–2013, the combined efforts of the French, the UN, and the Malian military have since made the area safer for researchers and participants. Third, since communal strife has pervaded life in central Mali in general and in Segou in particular, these respondents are very aware of the types of local disputes that occur and the conditions under which they escalate into violence. In sum, respondents from Bamako and Segou represent the full range of the Malian experience with communal disputes that international peacekeepers must address.

Vignette and Treatment

The survey began with a set of basic demographic questions and baseline questions about international actors that were identical for all respondents. Next, all respondents received a vignette describing a typical

Table 4.1 *Summary statistics and balance on demographic covariates between treatments*

	Mean			Difference			*P*-value		
	Control (C)	France (Fr)	UN	Fr-C	UN-C	Fr-UN	Fr-C	UN-C	Fr-UN
Age	34.54	35.40	36.24	0.86	1.70	−0.84	0.43	0.14	0.46
Female	0.33	0.30	0.33	−0.04	−0.01	−0.03	0.34	0.86	0.45
Children	2.84	2.98	2.93	0.14	0.08	0.05	0.61	0.75	0.85
Education (0–9)	3.00	2.98	3.18	−0.01	0.19	−0.20	0.96	0.45	0.43
Employment (0–3)	1.48	1.58	1.55	0.10	0.08	0.03	0.32	0.47	0.80
Victimized	0.30	0.32	0.32	0.01	0.01	0.00	0.72	0.72	1.00

communal dispute of the sort that happens on an almost daily basis all over the country. The vignette involved a land dispute over cattle between two ethnic groups:

Before the war, [family 1][6] herded their 80 cows on land which they owned. [Family 1] had bought the cows over many years and had owned this land for 35 years. In December 2013, [family 1] was forced to leave their land and cows under threat of violence from armed bandits. When they left, [family 2] seized the land and the cows that were left on the land. When [family 1] returned to their land earlier this year, [family 2] refused to give or sell the land or the cows back to them. Some of [family 1] now wants to take back their land and cows by threatening [family 2] using guns.

I randomly assigned respondents to either the control group or one of two treatment groups. Respondents were balanced on demographic covariates across treatments (see Table 4.1). Respondents in the control group received no further information. Those in the UN treatment group were told that two UN peacekeepers in the area discovered the dispute between the two families. Those in the France treatment group were told that two French peacekeepers came across the dispute. During the debrief, enumerators reiterated that the vignette was hypothetical.

After presenting respondents with the vignette and treatment, I asked them how likely they thought it was that violence would break out. Respondents could answer on a five-point scale, but for ease of interpretation I recoded the outcome as a binary variable in which "very likely" and "likely" are coded as 1 and all other responses as 0.

[6] I randomly varied the names of the families between four different names to avoid any bias due to specific association with a family name.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

Subnational Analysis of Communal Violence on the Burkina Faso–Mali Border

To analyze the relationship between peacekeepers' nationality and their ability to prevent communal violence, I use a time-series cross-sectional dataset of peacekeeping deployments to Mali. The primary analysis focuses on Mopti in central Mali, the region that experiences the highest levels of communal violence. The unit of analysis is the *commune*-month. *Communes* are the third-level administrative district in Mali (ADM3) and the smallest administrative district for which systematic data is available. The *commune* level offers more precision than the grid level, the spatial unit I use in the cross-national analysis in Chapter 3. I use the grid as the unit of analysis in the Mali–Burkina Faso comparison below, given its comparability across countries with different administrative units and government structures. I examine every month of available data from January 2012, the beginning of the conflict, to December 2019, the most recent month for which I have complete data on the necessary variables.

The dependent variable is a binary coding of the onset of communal violence from multiple event-based datasets. Since I conceptualize communal conflict primarily as violence that does not involve the state, I operationalize it in a similar manner: I code a *commune*-month as featuring communal conflict if violence occurs between two nonstate actors in that month in that *commune*. I use a binary coding of the outcome variable because I want to measure peacekeepers' ability to prevent the onset of any type of communal conflict.[7]

Given that communal conflicts can take many forms, it is unlikely that any single conflict event dataset will include all relevant types of violence. In the cross-national analysis described next I use Armed Conflict Location and Event Dataset (ACLED) data to capture as many communal violence events as possible. However, focusing on a single country allows me to gather more (and better-quality) data. Various event datasets have been created to record different types of activities. ACLED covers violent events as well as nonviolent actions such as the establishment of bases (Raleigh et al. 2010), the Global Terrorism Database (GTD) (LaFree and Dugan 2007) chronicles terrorist attacks, and the Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset (GED) contains attacks by groups in active civil conflicts that result in at least one fatality (Sundberg, Eck and Kreutz 2012). Moreover, some datasets may contain specific reporting inaccuracies that could generate biased estimates (Eck 2012).

[7] Since there were almost no cases of more than one communal conflict in a *commune* in a single month, alternative codings would not change the results of the analysis.

To compensate for the inaccuracies of each data source and to capture the broadest possible spectrum of communal violence, I combine multiple event datasets: ACLED, GED, the Social Conflict Analysis Database, and GTD (LaFree and Dugan 2007; Raleigh et al. 2010; Sundberg, Eck and Kreutz 2012; Salehyan et al. 2012). Since the datasets are created separately, they could include the same events; there is no universally agreed-upon system of identifiers that uniquely identifies events across datasets.

To minimize the risk of double counting events, I use a deterministic algorithm to identify groups of two or more events from the datasets that potentially refer to the same incidents (Donnay et al. 2019).[8] The algorithm flags all events that fall within the given date and distance window of other events, and then detects likely double reports using event features such as actor identity, event type, or confidence in the geocoding provided by the dataset. For instance, the algorithm flagged two events as a match and retained only one to avoid duplication: (1) one-sided violence perpetrated by the armed group Ansar Dine with the highest level of geographic precision (in UCDP GED) and (2) violence against civilians committed by Islamist militants with the second-highest level of geographic precision (in ACLED). It also produces a report of duplicated events, which allowed me to inspect the description of the source article in each dataset and further increase my confidence that the two observations describe the same event and that no event has been incorrectly eliminated. After running the program, I manually examined each observation to confirm the precision and accuracy of the data, ensuring that every observation captured a unique communal violence event.

Figure 4.4 maps all the communal violence events in the dataset. It indicates that most of the violence is clustered around Mopti, which demonstrates the region's instability. It also highlights the variability in the onset of communal violence across Mopti: Most conflicts are located in the eastern *communes*.

I leverage the fact that an international border between Mali and Burkina Faso separates an area of pervasive communal disputes to test the observable implications outlined earlier. The previous section explained how peacekeepers in Mali deploy to where they are needed most following an instrumental logic. Since they deploy in response to conflict, the areas where they are present will likely have the highest levels of violence. If we were to examine the relationship between peacekeeping deployment and violence, we might observe a null or even negative association. However, this does not mean that peacekeeping does not have a positive effect. Rather, it would indicate that we have not separated the effect

[8] I thank Rob Williams for his assistance in running this program.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

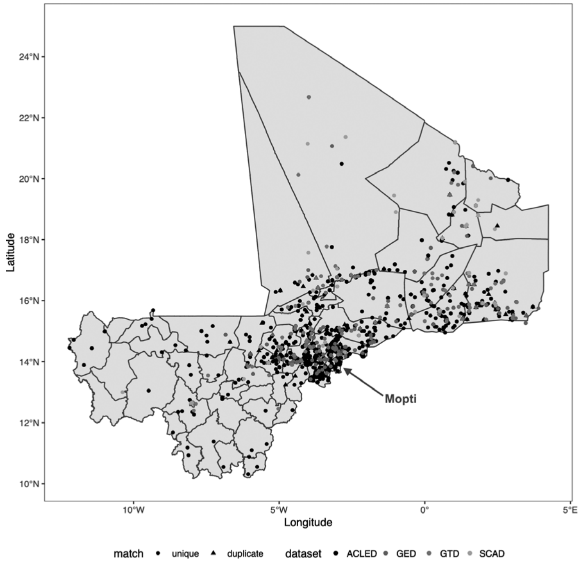


Figure 4.4 Communal violence in Mali, 2012–2020

of peacekeeping from the selection bias generated by their approach to deployment. It is thus especially important to develop an empirical approach that can identify a causal effect independent of any bias that may arise from these selection issues.

History of the Border

Mali and Burkina Faso share a border of approximately 1,300 km, a significant part of which has been historically disputed. The disagreement dates back to the colonial era when both countries were part of French West Africa. When Mali and Burkina Faso became independent in 1960, the border was not officially delineated (Naldi 1986). Disputes over western villages along the border as well as grazing rights and water sources ensued.

War over the border began in 1974. While the Organisation of African Unity soon negotiated an end to the violence, informal contestation of the border continued. Both countries brought the case to the International Court of Justice (ICJ) to investigate in 1983. While the case was pending, a three-day war broke out on December 25, 1985, killing thirty-five soldiers from Burkina Faso. Another ceasefire followed. The ICJ decided in 1986 to apply *uti possidetis* law to the case, meaning that the international community would recognize the borders that were in place at the end of the French colonial period; where the dividing line was unclear, the court simply divided each area in half. Both sides accepted the 1986 verdict and demonstrated a renewed initiative to maintain peace (Leigh 1987).

The court ruling settled the dispute between the two countries and formally established an internationally recognized border. However, in practice the border remained permeable to both civilians and armed groups. Numerous local militants have recently exploited this fact to spread from Mali to northern Burkina Faso. The border areas are thus very unstable, and armed groups are prominent. While people and goods continue to move freely across the border, UN peacekeepers cannot because they are bound to uphold international law and the 1986 ICJ ruling. I leverage this natural experiment in the analysis later.

Identification Strategy

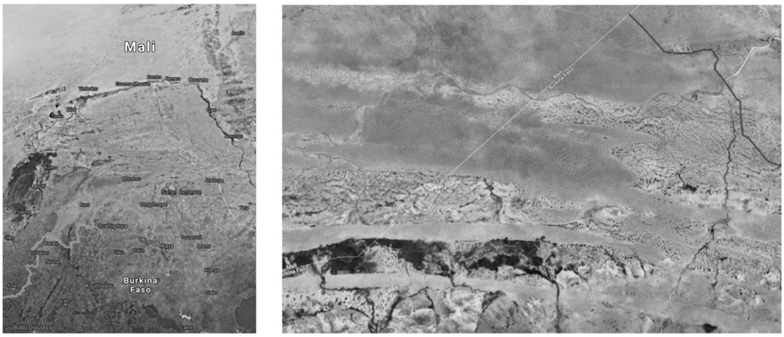
I use a geographic regression discontinuity design (GRDD) to identify the effect of peacekeeping patrols.[9] The GRDD measures the local average treatment effect at a geographic boundary that splits observations into treated and control areas as-if randomly. Implementing a GRDD requires restricting observations close to the boundary and measuring a running variable that indicates each village's distance to the boundary. Since peacekeepers cannot cross into Burkina Faso, comparing grids on either side of the border allows us to measure the impact of peacekeepers on the Mali side.

The border between the two countries is not clearly marked, and there is little to indicate the border on the ground (Figure 4.5). Recent studies suggest that the observable and unobservable confounding variables that might play a role in violence are the same on both sides. According to ongoing work by Michael Kenwick and Beth Simmons, there are only two developed border crossings between the two countries, and the border as a whole is around the 10th percentile for fortification in the world.[10] I restrict the sample to the territory around the border between those two crossings, which is not distinguishable on the ground.

The key concern with GRDDs is compound treatments – that is, that more than one treatment may affect the outcome. In this context, the border may create multiple treatments. For example, if a particular benefit is only available to those on the Burkina Faso side of the border, that would create a compound treatment. Because the border is indistinguishable in practice, I argue that whatever differences may exist in the countries generally, the areas I study do not reflect them.

[9] See Keele and Titiunik (2015). For another applied example, see Henn (2023).
[10] Author correspondence with Kenwick.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press



(a) Satellite image, border (b) Satellite image, part of the border

Figure 4.5 Burkina Faso–Mali border
Attribution notes: Imagery ©2024 TerraMetrics, Map data ©2024 Google Maps.

Another major concern for GRDDs is spillover: The treatment being measured can influence units that are not assigned to the treatment. The current case has two potential forms of spillovers. The first is that peacekeepers – intentionally or, more likely, unintentionally – cross the border into Burkina Faso. However, interviews conducted with UN officials in the field suggest that this is highly unlikely because this would lead to sanction from a commanding officer or expulsion from the mission. Peacekeepers use GPS devices to determine their patrol routes and keep to the Malian side.[11]

A second possible source of spillover is that people in Burkina Faso may believe that peacekeepers would intervene on their behalf. This is also unlikely, given that they would never have experienced any intervention on their behalf from peacekeepers. Even if they believed it would happen when the peacekeepers first deployed in 2013, the fact that it has not occurred would change their beliefs over time. While I address spillovers more directly in the robustness checks, in general both forms of spillovers across the boundary should downward bias my results by making Burkina Faso more, not less, similar to Mali.

The geographic unit of analysis is a grid cell approximately 10 km $\times$ 10 km, and the temporal unit is the year-month (e.g., from January 2014 as the first unit to December 2020 as the last unit). I use grids rather than villages because communal violence sometimes breaks out in unmarked rural areas in this part of Mali and Burkina Faso. To create the treatment variable, I restrict the sample to grids within 100 km of the border area.

[11] Author interview with UN officials in MINUSMA.

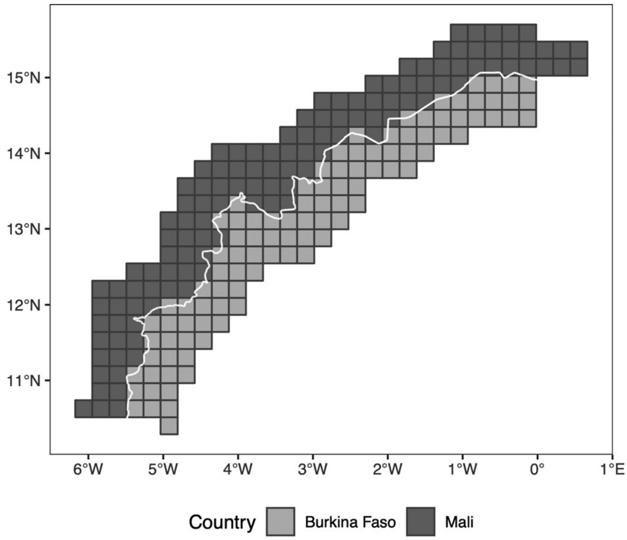


Figure 4.6 Grid sample
(Source: Author's own illustration)

On the Mali side of the border, I further limited the sample to the stretch of the border that is located in the Mopti region to provide a conservative estimate of the treatment and to more cleanly identify the scope of peacekeeping. Since communal violence has been particularly severe in Mopti, patrols have focused on local-level peacekeeping rather than other peacebuilding tasks in this area. For example, in the northern regions of the country peacekeeping troops more frequently engage armed groups directly and are involved in other peacebuilding dimensions.

Figure 4.6 displays the sample of all grid cells within 100 km of the border. Grid cells on the Mali side are assigned to treatment (dark gray) and those on the Burkina Faso side are assigned to control (light gray). Grid cells on the border are assigned to the country that constitutes more of their area.

To bolster the validity of the estimates, I further control for four relevant factors that may impact the level of violence in a given area. First, I control for the level of nighttime light emissions measured by satellite in 2011.[12] Nighttime luminosity is a common proxy for economic activity,

[12] I use data from 2011 to avoid potential posttreatment bias because the conflict in Mali began in 2012.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

and areas with different levels of economic development may experience different rates of violence.[13]

Second, I control for the size of the Fulani population in each grid cell. As described in Chapter 4, the Fulani are one of Mali's major ethnic groups. Much communal violence in central Mali involves Fulani, so controlling for their presence is important for obtaining valid estimates of the effect of UN peacekeeping. This data is drawn from the Spatially Interpolated Data on Ethnicity dataset (Müller-Crepon and Hunziker 2018).

Such a binary treatment variable, however, disregards potentially important variation: It treats treatment grids identically even though they may not receive the same amount of patrolling. For this reason, I also control for two additional factors – the number of peacekeepers at the nearest UN base and the distance from the center of the grid to the nearest UN base.

Peacekeeping Deployments in Mali

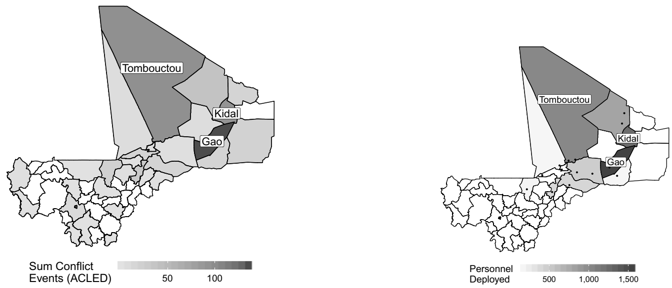
I use data from my RADPKO dataset to code the independent variable – the presence of peacekeepers.[14] I first use the data to investigate where and when UN peacekeepers deployed within Mali. The broad patterns of deployment illustrate that peacekeepers follow what Ruggeri, Dorussen and Gizelis (2017) call an "instrumental logic" rather than a "logic of convenience." That is, they deploy to help resolve conflicts rather than to safe areas where the risk of attack is low. Nor do I find evidence of a large temporal lag between violence and deployment. In other words, UN peacekeepers go where it is dangerous (while it is still dangerous), not where they are at lowest risk of being killed.

For instance, violence in Mali before the UN deployment was particularly severe in three second-level administrative districts (ADM2s) – Tomboctou, Kidal, and Gao. The UN deployed the most peacekeepers to those three districts both in the first month of its deployment and over the duration of the mission (see Figure 4.7). It also deployed a smaller yet significant number of peacekeepers to the next three most violent districts – Tessalit, Ménaka, and Douentza – suggesting the number of peacekeepers it sends is proportional to security threats.

As more specific evidence of these reactive deployments, consider the increase in peacekeeper deployment to the Tomboctou region in Mali from 1,250 peacekeepers in May 2015 to 1,600 in June to 2,200 in July (see Figure 4.8). Over the same period, there were six conflict events

[13] This data comes from the Defense Meteorological Satellite Program Operational Linescan System (Elvidge et al. 2009).

[14] I describe this dataset in greater detail later in the chapter.



(a) Violence in Mali before UN PKO, by region (b) UN personnel, by region

Figure 4.7 UN deployment to Mali (MINUSMA)

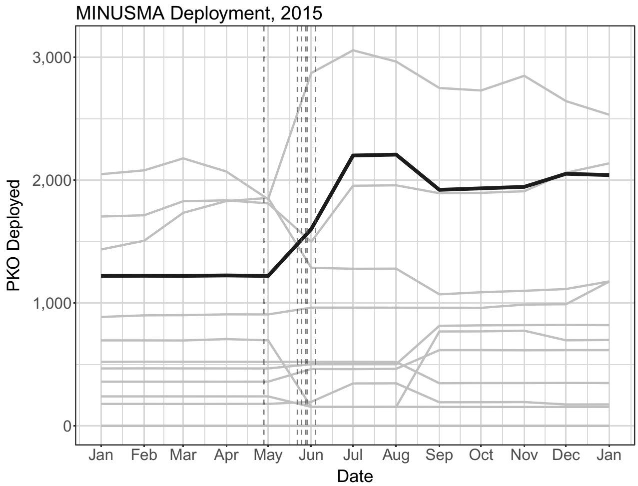


Figure 4.8 MINUSMA reactive deployment, Tombouctou, 2015
Note: Vertical red lines mark the dates of violent events involving MINUSMA troops, as recorded in the ACLED database. Faded gray lines track the deployment of UN peacekeepers to other treated ADM2 units in Mali; the single blue line tracks the number of peacekeepers deployed in Tombouctou (*cercle* administrative level).

in the region, five involving UN peacekeepers. In reports published during this period, UN officials expressed concern about these attacks and formally requested more troops in the country and in Tombouctou specifically. These reports suggest that these UN troop increases in Tombouctou reflect an explicit concern about the stability of the Malian state.[15]

[15] Report of the Secretary-General on the Situation in Mali, June 11, 2015; Report of the Secretary-General on the Situation in Mali, September 22, 2015.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

Subnational Analysis of Peacekeeping in Central Mali

Next, I examine the entire region of Mopti, which contains the Mali side of the border region examined in the GRDD. Rather than comparing UN peacekeeping in a country with a PKO (Mali) to a country without one (Burkina Faso), I compare the actions of some UN peacekeepers from one country to those from another country *within Mali*.

UN Peacekeeping in Mopti

Mopti is relatively unusual among peacekeeping areas in Mali because peacekeeping patrols have primarily been undertaken by troops from Togo and Senegal, two francophone countries in West Africa. Troops from other countries primarily perform technical duties such as detecting mines and improvised explosive devices.

Although both groups of peacekeepers speak the same language (French) and are from the same region of the world (West Africa), Togolese peacekeepers differ in important ways from those from Senegal. The ethnic makeup of Togo is different from Mali and, in particular, Mopti. Whereas the Fulani are a dominant group in the Mopti region and number almost 3 million inhabitants across Mali, there are fewer than 1,000 Fulani in Togo. Malians are also unlikely to perceive Togolese peacekeepers as religiously biased, since only 14 percent of the Togolese population is Muslim. Thus the identity cleavages in this part of Mali are irrelevant to Togo.

Senegal, however, has similar identity cleavages to the Mopti region. The Fulani are an influential minority group in Senegal, and enjoy national-level political power. The president of Senegal from 2012 to 2024, Macky Sall, is one of the most powerful Fulani in West Africa. As a result, non-Fulani in Mopti are likely to perceive Senegalese peacekeepers as biased in favor of the Fulani in Mali. Given the importance and salience of Fulani/non-Fulani relations in central Mali, localized peace enforcement theory predicts that these perceptions will shape Senegalese peacekeepers' ability to limit communal violence.

Specifically, localized peace enforcement theory predicts that Senegalese peacekeepers will fail to limit the escalation of disputes as they cannot signal their lack of bias to local populations on both sides. By contrast, as long as they have enough troops deployed to a given locality, Togolese peacekeepers will be able to signal their lack of bias and prevent disputes from escalating. Other theories would predict that Senegal would perform as well as (if not better than) Togo. Peacekeepers from both countries speak French, suggesting they are able to communicate with the local population. Senegal is also more culturally proximate to

Mali than Togo and to probably any other country with peacekeepers in Mali with the exception of Burkina Faso and maybe Niger. Finally, there are more than enough peacekeepers from both countries to deter violence. Figure 4.9 graphs the number of Togolese troops deployed to Mopti over time. On average, there were 200 to 300 troops from Togo in each locality. Although there were no Senegalese peacekeepers in Mopti from late 2014 to late 2017, the number rose sharply in 2018 and 2019.

This comparison allows me to control for two alternative explanations of local-level peacekeeping success. According to the *capacity explanation*, as the number of peacekeepers deployed to Mopti increases, communal violence should decrease, regardless of the peacekeepers' nationalities (i.e., there should be no difference between Togolese and Senegalese peacekeepers). Similarly, the *biased peacekeeping explanation* would expect peacekeepers from both countries to gather information about disputes equally effectively and, as a result, to contain disputes equally well. According to this explanation, both groups of peacekeepers should be especially successful at this task given their relative cultural proximity to Mali. Senegalese peacekeepers should arguably be more effective, given that Senegal is culturally more similar to Mali than Togo is.

Data and Estimation Strategy

The explanatory variables are counts of UN peacekeeping troops from Togo and Senegal. I use data from the RADPKO dataset to measure the number of peacekeepers deployed from Togo and Senegal in a *commune*-month in Mopti. RADPKO collects monthly data from publicly available UN reports on the location of UN peacekeeping deployments, divided by every peacekeeping-contributing country.[16] As Figure 4.9 shows, these two countries account for the vast majority of troop deployments in the area. Although there are some other contributors – such as a small number of Cambodian and Egyptian peacekeepers deployed in late 2019 – these troops perform primarily technical duties such as mine and improvised explosive device detection.[17]

I estimate the association between peacekeepers and the onset of violence using logistic regression models since the outcome is a binary variable. To mitigate omitted variable bias, I adjust for a set of potential covariates that might be associated with both deployment patterns

[16] For more on the coding protocol, see the discussion of the cross-national research design later in the chapter. The monthly data is a unique feature of the RADPKO data, making it more suitable for subnational, time-series analyses than other UN data. See Bove, Salvatore and Elia (2021) for a recent application using RADPKO.

[17] All models include counts of Cambodian and Egyptian peacekeepers as controls to adjust for these deployment patterns.

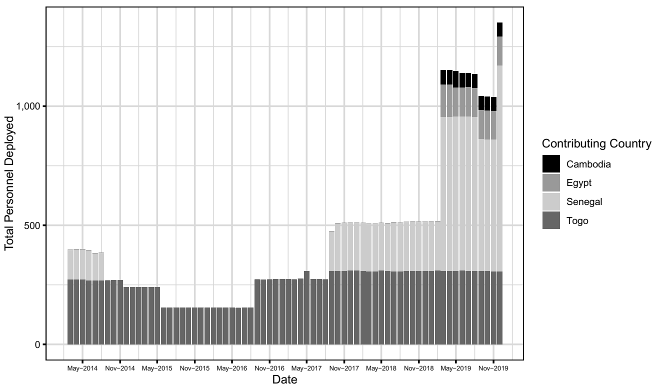


Figure 4.9 Average Togo and Senegal troop deployment

and the onset of conflict. Given the importance of geographic distance, infrastructure, and terrain to the effectiveness of peacekeeping troops (Ruggeri, Dorussen and Gizelis 2017), I include measures of the distance between the *commune* and the nearest peacekeeping base, the level of development in the *commune* as measured by nighttime luminosity, and the ruggedness of the terrain as measured by a *commune's* average elevation multiplied by the slope. The models also account for temporal dependence between units using a cubic polynomial (Carter and Signorino 2010). Finally, I add fixed effects at the *cercle* level, the second-level administrative district (ADM2), to account for potential differences between units that the covariates may not capture. I include robust standard errors clustered at the *cercle* level.

Before proceeding, I discuss three potential threats to causal inference and my strategy for mitigating them. First, the strategic nature of UN peacekeeping deployment may introduce selection bias because UN peacekeepers do not patrol conflict settings at random. For the purposes of my analysis, deployment patterns are primarily a concern if peacekeepers go to safe areas rather than where they are needed most, but as I have described, peacekeepers follow what Ruggeri, Dorussen and Gizelis (2018) call an "instrumental logic" rather than a "logic of convenience," and they do so without a large temporal lag. Additionally, as discussed earlier, I control for the geographic characteristics of the *commune* to account for selection bias that may arise from peacekeepers' decisions to avoid areas that are difficult to reach. Finally, given that I compare peacekeepers from two West African countries under the same UN command structure, there is no reason to expect that concerns related to selection could explain differences between the two.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

Table 4.2 *Summary of observable implications from Hypothesis 3 of localized peace enforcement theory*

	Main Analysis	Placebo Test
	Mopti	*Kidal*
Togolese peacekeepers	Perceived as impartial, lower likelihood of violence	Perceived as impartial, similar likelihood of violence
Senegalese peacekeepers	Perceived as biased, higher likelihood of violence	Perceived as impartial, similar likelihood of violence

Note: Each cell predicts whether Malians will perceive peacekeepers from the country of origin (Togo or Senegal) as impartial in the Malian region of deployment (Mopti or Kidal) and the relative likelihood of communal violence.

A second potential issue is that the dataset does not capture all communal disputes that occur *or* whether they become violent. I mitigate this concern by examining a cross-section of data recording violent events in localities in one specific region of a country in which communal disputes are pervasive. Theoretically, each *commune* in Mopti should have similar numbers of disputes and a similar likelihood of communal disputes turning violent, all else equal. By measuring whether communal conflict occurred at a specific time and place, the data reflects whether these localities with similar dispute propensities became violent.

Finally, a critical concern is whether the results can be attributed to the proposed mechanism related to perceptions of impartiality or another mechanism altogether. I address this issue using a multifaceted design-based approach that compares potential observable implications from the theory to alternative mechanisms (see Table 4.2). According to the logic of my theory, the deployment of Togolese (but not Senegalese) peacekeepers should be negatively associated with the onset of communal violence. Moreover, this relationship should be localized to the region of Mopti since ethnic power relations there resemble those in Senegal but not Togo. I test this implication by conducting a placebo test using data from another region of Mali, Kidal, that has entirely different ethnic power relations than either Senegal or Togo. Given the lack of ethnic connection to the domestic populations, Malians living in Kidal will likely perceive Togolese and Senegalese peacekeepers as similarly impartial. Consequently, my theory predicts that peacekeepers from both contributing countries will be equally effective.

Cross-national Analysis

Measuring Peacekeeping Deployment

Hypothesis 3 predicts that deploying UN peacekeepers to a community reduces the outbreak of communal violence. To measure the presence of such patrols, I need to operationalize peacekeeping presence to a high degree of geographic and temporal certainty since deployments in response to violence shift rapidly. To do so, I created a comprehensive dataset of UN PKOs conducted at the local level in collaboration with Patrick Hunnicutt.

This information was previously not available. A special issue of *International Peacekeeping* edited by Govinda Clayton in 2017 described the state of the art with regard to peacekeeping data (Clayton et al. 2017). Han Dorussen and Andrea Ruggeri discuss the PKOLED and PKODEP datasets in this issue, which jointly identify the time and subnational location of UN peacekeeping deployments in Africa from 1989 to 2006 (Dorussen and Ruggeri 2017). Along with Theodora-Ismene Gizelis, they use this data to analyze the effectiveness of UN PKOs at the local level during this period (Ruggeri, Dorussen and Gizelis 2017). Hanne Fjelde, Lisa Hultman, and Desireé Nilsson collected data on the time and subnational location of UN peacekeepers deployed to prevent civilian victimization in nine African countries from 2000 to 2011 (Fjelde, Hultman and Nilsson 2019), which they later extended to 2014 in work with Deniz Cil (Cil et al. 2019).

Hunnicutt and I used primary documents collected from the UN to construct the RADPKO dataset, which contains the location and number of UN peacekeeping troops, UN police, and UN military observers deployed to every new UN PKO authorized under a "robust" Chapter VII mandate from 1999 through 2019.[18] RADPKO also records the gender, specialization, and nationality of all deployed personnel. As Table 2.1 shows, all Chapter VII operations approved during this time were in Africa, hence the name RADPKO. We do not include the deployment of troops from regional organizations such as the Economic Community of West African States or the African Union or single countries like France or the United Kingdom.

Our dataset represents an improvement in the quality of data on local-level peacekeeping for three reasons. First, our data solely and entirely encompasses the scope of UN PKOs that have the authority to patrol at the local level. Second, it offers precise estimates of all UN

[18] This dataset is published as Hunnicutt and Nomikos (2020). For access to the data, see the online interface for the RADPKO data on the website of the Data-driven Analysis of Peace Project, https://dapp-lab.org.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

peacekeeping deployments in Africa. We use primary data on country-level force contributions sourced directly from the UN Department of Peace Operations (DPO), which is updated monthly. In a third improvement, our dataset includes previously unavailable fine-grained information on peacekeepers disaggregated by type, nationality, and gender. With an eye toward replication, we worked to confirm, refine, and expand upon existing data. This section describes these efforts in detail.

Data Collection For every Chapter VII peacekeeping mission deployed to sub-Saharan Africa from 1999 to 2019, RADPKO records estimates of the location, date, and count of UN peacekeeping personnel deployed by each contributing country, disaggregated by personnel type (e.g., police, military observer, or troops) and gender. We collected this information from two types of archival documents from the DPO.

First, we use deployment maps, which are periodically available in mission reports to the Secretary-General, to identify the complete set of active peacekeeping bases and generate a count of peacekeeping units per nationality deployed at each base in a given month. The UN issues reports from the Secretary-General to the Security Council every three months for each operation. Each report contains a deployment map that pinpoints the geographic location of all UN bases. To code a variable in the time-series data to indicate the presence of a UN base, we went through each report for every PKO and pinpointed where the UN placed its bases. If the location remained the same from one report to the next, we assumed it had not moved. If it did move, we assumed it moved at the beginning of the new period of reporting.

Second, we draw on monthly DPO deployment reports to generate contributing-country-level counts of peacekeeping personnel by type and gender for all active Chapter VII missions. These publicly available reports provide data on the number and type of peacekeeping troops. Although they do not detail the exact number of UN peacekeeping troops/bases, they do list the identity of UN peacekeeping battalions at each base and the number of peacekeeping troops contributed by each country. This allows me to match the size of peacekeeping contributions to the battalions at a given base location gathered from the Secretary-General reports.

Figure 4.10 graphs the RADPKO data, which illustrates that across the 12[19] missions in the dataset, there is substantial variation over time in the deployment of peacekeepers. This variation underscores the importance of examining peacekeepers' effectiveness using dynamic data that varies over time.

[19] I treat MONUC and MONUSCO as one mission.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

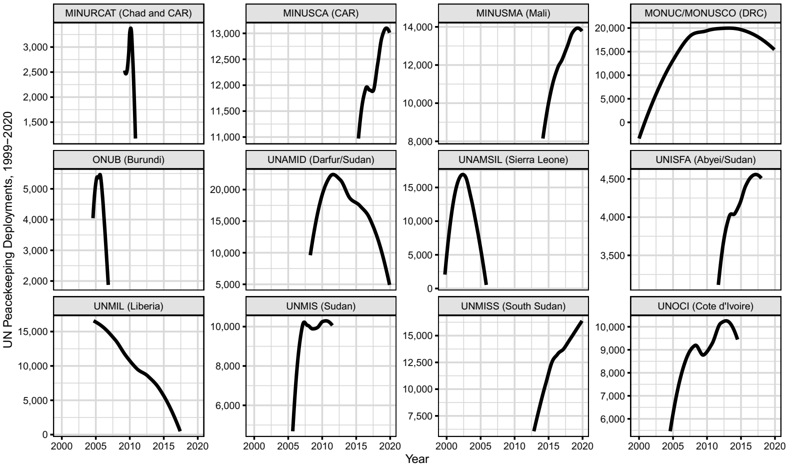


Figure 4.10 UN peacekeeping deployments to Chapter VII-authorized missions in Africa, 1999–2020
Note: For ease of interpretation, the Y-axis is scaled differently for each mission. Source: RADPKO.

Contribution and Comparison to Other Data

RADPKO offers uniquely precise estimates of peacekeeping contributions at the local level for at least two reasons. First, relying exclusively on map symbology, as other datasets do, to estimate the number of personnel in each unit ignores the likelihood that unit size may vary by operation. For instance, a company of Bangladeshi troops deployed in Sierra Leone may be significantly smaller than a company of Bangladeshi troops deployed in Mali, given the different operational constraints associated with each context. Second, the symbology-based estimation technique risks generating incorrect estimates of each contributing country's unit size whenever the actual unit size does not perfectly match symbology-based standards for unit size (10 troops per squad, 35 troops per platoon, 150 troops per company, and 650 troops per battalion). There is no way to determine when these standards may lead to over or underestimates of actual deployment levels, and no qualitative evidence regarding whether Chapter VII missions strictly adhere to these standards when deploying troops. Audits of ongoing peacekeeping missions suggest that deployed units are commonly understaffed, which implies a constant but unmeasurable degree of uncertainty associated with symbology-based estimates of personnel. Our data instead uses exact counts of deployed peacekeepers per contributing country, type, and gender to estimate peacekeeper force size at the subnational level.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

Additionally, our data better captures temporal variation in the UN's subnational deployment of peacekeepers. No other dataset uses reports on *monthly* force contributions to estimate the number of peacekeepers deployed from each active contributing country for each mission. The DPO does not regularly publish publicly accessible mission reports from the office of the Secretary-General that contain updated deployment maps. Relying on deployment maps alone to estimate deployment statistics therefore likely overlooks monthly changes in peacekeeping deployments that may occur in response to mission-specific needs. Thus other available data sources cannot capture variation in the number of peacekeepers deployed in months in which a new deployment map is not published. Our data can capture variation in peacekeeping force contributions at the local level for the many months for which deployment maps are not available – a particularly important advantage given the dynamic nature of Chapter VII peacekeeping deployments.

We believe RADPKO contains the most comprehensive available data on local-level peacekeeping. But it is not without limitations. Notably, we are missing data for each mission in our sample: We are only able to include updated information on the location and composition of deployed peacekeeping units for months in which missions publish a report of the Secretary-General containing a deployment map. Thus for each mission, we are missing data from the months prior to the release of the first such map and the months between the publication of each subsequent map. All available data on the subnational deployment of peacekeeping forces faces the same limitations, since previous data collection efforts relied solely on these reports to estimate subnational deployment statistics.

Unfortunately, we cannot resolve either point of missingness using publicly available UN DPO documents. Instead, we make two assumptions. The first is that a mission's first publicly available deployment map represents the location and composition of peacekeeper deployment from its establishment. Second, we assume that a mission's location and composition of peacekeepers remain static between the publication of Secretary-General reports with updated deployment maps.

Measuring Capacity: Base Presence, Peacekeeping Troops, and Distance In line with previous work (Ruggeri, Dorussen and Gizelis 2017; Fjelde, Hultman and Nilsson 2019), I divide each country into 0.5×0.5 decimal degree grids with the month as the temporal unit of analysis; 0.5 decimal degrees are about 50 km at the equator, meaning that each grid is roughly 2,500 km^2 in size. I then aggregate peacekeeping deployments within each grid cell for each month. Spatial grid cells are a unit of analysis that is not endogenous to conflict processes, reduces the degree of measurement error present in the dependent variable, and recognizes the conflict's spatial and temporal dynamics. Since communal disputes are not defined by the boundaries of a particular administrative district, it is more accurate to use a geographic measure to bound a unit of observation. The grid version of the dataset has the added advantage of being readily mergeable with all data available from the Peace Research Institute Oslo's UCDP (Tollefsen, Strand and Buhaug 2012), which I use to collect the georeferenced covariates discussed later.

According to localized peace enforcement theory (presented in Chapter 3), peacekeepers' ability to prevent communal violence rises as a function of the cost they can impose on any party that chooses to defect from cooperation. Thankfully, the RADPKO data offers several options for measuring local peacekeeping capacity, the most straightforward of which is a binary indicator of the presence of a UN peacekeeping base in a given grid in a given month. However, this operationalization underestimates peacekeepers' effectiveness in two ways. First, it assumes that all peacekeeping units are equally effective. Since this is not the case, I add the count of peacekeepers from RADPKO. For each grid-month, I record the number of peacekeepers deployed and scale this number by 1,000 for ease of interpretation: One unit for each independent variable corresponds to 1,000 UN personnel. Second, this approach assumes that peacekeeping effectiveness is contained within these grid areas, yet some peacekeepers patrol far from the capital or their base. To account for this possibility, I estimate the minimum linear distance between the center of a grid cell and the nearest UN PKO base.

Since we seek to estimate how deploying peacekeepers affects the likelihood that violence will break out in the future, we must modify the data accordingly. I therefore introduce a lag for the peacekeeping variables so that in month t, we are observing the relationship between peacekeepers deployed at time $t - 1$ and the onset of violence at t.

Measuring Perceptions of Impartiality

Localized peace enforcement theory predicts that peacekeepers are more likely to achieve their objective as the perceived probability that they will leverage capacity to punish both parties increases. When the perceived probability that UN peacekeepers will punish both parties is equal and high, local populations consider the UN peacekeeping contingent to be relatively impartial. The primary implication of Hypothesis 3 is that peacekeepers successfully prevent communal disputes from escalating when domestic groups perceive them as impartial. Although it is extremely challenging to measure cross-national perceptions of impartiality using observational data, I use peacekeepers'

Table 4.3 *Average number of peacekeepers deployed, by mission*

Mission	Grids	All	Western	Regional	African Nonregional	All Impartial
MINURCAT	493	2,677	943	0	1,017	1,734
MINUSCA	257	12,266	233	2,508	5,180	9,524
MINUSMA	503	12,080	1,069	6,504	1,452	4,507
MONUC	850	11,818	154	0	3,502	11,664
MONUSCO	850	188,42	327	0	5,057	18,515
ONUB	18	4,514	16	0	2,480	4,498
UNAMID	217	16,852	16	722	10,515	16,113
UNAMSIL	40	11,019	591	3,441	1,565	6,987
UNISFA	11	4,164	4	0	4,127	4,160
UNMIL	51	9,963	507	2,615	1,146	6,841
UNMIS	936	9,644	394	617	1,055	8,633
UNMISS	255	12,434	311	475	4,209	11,647
UNOCI	136	8,994	150	2,395	1,112	6,449

nationality as a heuristic for the degree of perceived impartiality. In Part II of the book (Chapters 5–7), I present micro-level evidence that supports the mechanisms underlying these patterns.

Although domestic populations are likely to view all UN personnel as less biased than peacekeepers from a single country on average, we should also observe variation in perceptions of peacekeepers from different countries – especially those from Western countries (i.e., Europe, North America, Australia, or New Zealand) versus countries in the region. While peacekeepers' country of origin does not exactly capture perceptions of bias, it is the best possible proxy at such a coarse level of analysis. Readers should be aware of the limitations of this data and interpret it *in conjunction with* the analysis in Part II.

My argument is related to what Bove, Ruffa, and Ruggeri (2020) call indices of distance, which they measure in terms of geography, language, religion, political institutions, and economy. They find that cultural distance is associated with negative peacekeeping outcomes, and attribute this finding to the inability of peacekeepers from culturally distant countries to understand and interact with local populations, which aligns with the findings of qualitative researchers (Pouligny 2006; Autesserre 2010). By contrast, I argue that these differences are related to perceptions of bias dating back to colonialism (see Table 4.3).

Perceptions of Western peacekeeper bias arise in sub-Saharan Africa for at least three reasons. First, local populations might think they are troops from a former colonial power. Given that much of the effectiveness of localized peace enforcement relies upon perceptions, this simple confusion can have important effects. Further, it is not clear that an individual resident of a conflict or postconflict setting will be able to distinguish between a UN peacekeeper from a traditional colonial power like France or Britain and one from another Western country like Norway or Sweden, even if those differences may seem obvious to residents of these countries.

The second reason for perceptions of Western peacekeeper bias is that in several settings, a former colonial power has intervened alongside the UN (e.g., Sierra Leone, Mali, Côte d'Ivoire, and Central African Republic [CAR]). These interventions typically cause more civilian casualties than UN peace operations and, by definition, lack the same level of multilateral support. This makes confusion between white UN peacekeepers and colonial interveners even more likely. For instance, in the CAR, domestic and international observers both accused UN peacekeepers of the systematic sexual abuse of minors. However, it was not UN peacekeepers but French soldiers deployed to the CAR at the same time – independently of the UN – that committed the abuse (Howard 2019*b*).

Third, domestic groups may believe UN personnel from a former colonial power will be biased in favor of their preferred ethnic groups, despite their UN affiliation. The presence of UN personnel from former colonial powers may trigger feelings of "being imposed upon" (Talentino 2007). Local populations may not explicitly believe that UN personnel have colonial-era biases, but the presence of peacekeepers from colonial powers may trigger implicit perceptions about them. The relatively unknown affiliation of non-Western UN personnel is an important advantage in this context: Locals will have relatively little information about the biases of UN personnel from countries like India or Uruguay.

I also expect domestic populations to be more likely to perceive peacekeepers from countries with relatively similar identity cleavages as biased. Locals may have more information about the biases of UN personnel from countries they know more about, including neighboring countries. Such biases could work against those peacekeepers. However, as I discuss in the study of peacekeepers from Togo and Senegal in Mali in Chapter 7, this may vary depending on the contributing county. To operationalize similar identity cleavages, I take the count of peacekeepers deployed from contributing countries in the same region as the country of deployment. For each setting, these countries are likely to be the most socioculturally proximate to social groups in Mali. Bove, Ruffa and Ruggeri (2020) measure cultural distance as geographic contiguity. I broaden their approach to capture countries from the entire region because residents of a country may have similar identity cleavages as those from a country that does not necessarily share a border. I do not use any of their other distance measures because they are operationalized to range from 0 (personnel are culturally identical to the local population) to 1 (personnel are as culturally distant from the local population as possible), which prevents me from calculating and imputing peacekeepers' cultural distance in grid-months in the sample in which no personnel are deployed.

Measuring the Onset of Communal Violence

I rely on ACLED data (Raleigh et al. 2010) to operationalize outbreaks of communal violence. This dataset includes reported information on violent events coded by the perpetrator, type of event, date, and geographic location, which allows me to pinpoint communal violence events. I aggregate this data at the grid-month level and code a binary dummy variable that indicates whether communal violence broke out in that grid in a given month. I code the onset of violence based on what ACLED calls "identity militias":

> ACLED includes a broad category of "identity militias" that signifies armed and violent groups organized around a collective, common feature including community, ethnicity, region, religion or, in exceptional cases, livelihood. Therefore, for ACLED's purposes, identity militias include those reported as "tribal", "communal", "ethnic", "local", "clan", and "religious" and "caste" militias. *Events involving "identity militias" are often referred to as "communal violence" as these violent groups often act locally, in the pursuance of local goals, resources, power, security, and retribution* (ACLED codebook, emphasis mine).

The event-based data on violence used in this chapter likely undercounts the onset of violence. There are other possible datasets I could use to complement or replace ACLED, such as the UCDP GED nonstate actor dataset (Sundberg and Melander 2013) or the Social Conflict Analysis Database (Salehyan et al. 2012). However, they contain even fewer observations of communal violence than ACLED. Thus, although I acknowledge Eck's evaluation that UCDP's "geocoding and precision information is far superior to ACLED's" (2012, p. 137), I trade that precision to gain exponentially more observations. In the subnational analysis of communal violence in Mali in Chapter 7, I combine events from all of these databases and examine each observation manually, which provides precision *and* comprehensive coverage. This is unfortunately not feasible for the cross-national data.

Another methodological concern associated with using the ACLED data is that it might generate measurement error. It is difficult to capture all escalations resulting from every potential communal dispute. However, any such potential measurement error is likely to be nonsystematic, meaning that measurement error of the dependent variable simply produces statistical noise that is not correlated with the independent variables. For this reason, any measurement error will not bias the results of the statistical analysis. As these methodological shortcomings make clear, a deeper analysis of a single mission, which I offer in Part II of the book, is needed to thoroughly evaluate the hypotheses derived from the theory.

The nature of variation in the dependent variable illustrates why a subnational time-series approach is needed. For instance, there is *geographic* variation in the data. Consider the difference between and within cases visualized in Figure 4.11. Chad has had very few instances of communal violence, while the Democratic Republic of the Congo (DRC), Mali, Sudan, and South Sudan have had quite a few. Moreover, some countries, like South Sudan and the CAR, have had a fairly equal distribution of violence across grids. In other cases, like Sudan and Mali, communal violence is extremely localized: Certain areas are hubs of communal violence while others have experienced no violence at all. Analyses using data aggregated at the country level might adequately explain variation in UN effectiveness in cases like Chad or even South Sudan and the CAR. However, strictly country-level data will miss the spatial spread of violence in cases like the DRC and Mali.

The data also exhibits *temporal* variation. Figure 4.12 graphs the number of communal violent events during the UN PKO in the three countries with the most communal violence – the DRC, Mali, and Sudan/South Sudan. As I explain in Chapter 7, there is considerable variation over time even within countries, which collapsing the data into a single cross-section would miss. Sometimes the pattern seems fairly obvious, such as the rise in communal violence in Mali beginning in 2016. In other cases, the pattern is more subtle, as in the DRC where violence has steadily increased since 2007 after decreasing between 2000 and 2007. In yet other cases, the pattern is not at all obvious, as in Sudan, where violence has repeatedly fluctuated over time. And, of course, there is subnational variation over time as well. For all of these reasons, I use a cross-national dataset that varies within countries and over time.

Establishing Causality and Adjusting for Selection Bias

Selection bias is a key methodological concern associated with observational studies. Selection effects are problematic because an endogenous selection process can produce biased coefficient estimates. Changes in the local deployment of peacekeepers may be correlated with observable and unobservable factors that independently explain why armed groups target peacekeepers. One major threat to inference is that shifts in the broader conflict environment may explain both where peacekeepers are deployed *and* why communal violence breaks out. If UN peacekeeping

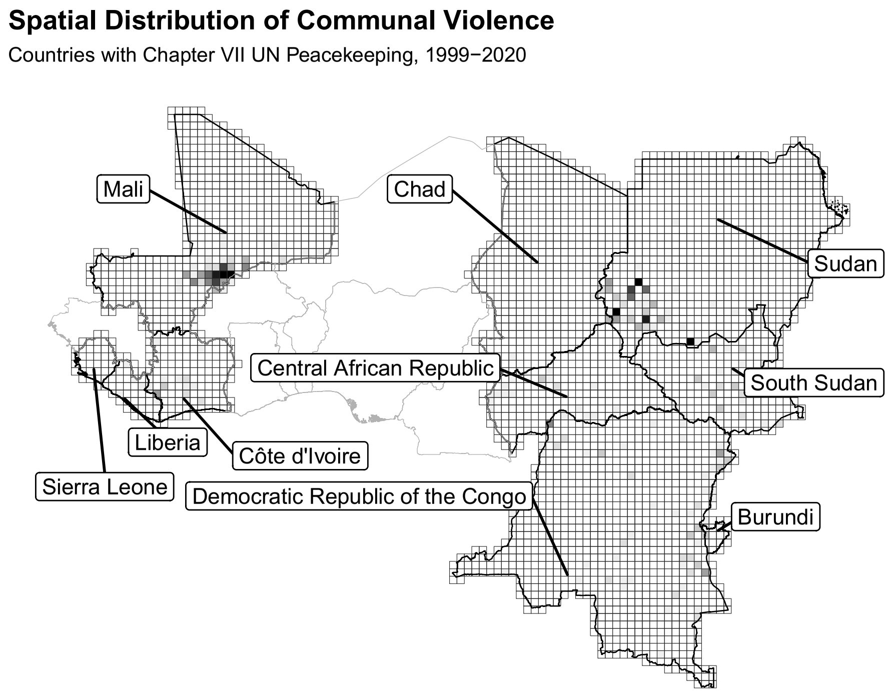


Figure 4.11 Total communal violence events
Note: 0.5 degree × 0.5 degree grid across ten countries in the sample. All shaded grids include at least one instance of communal violence. Darker shades indicate more violence from 0 to 14.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

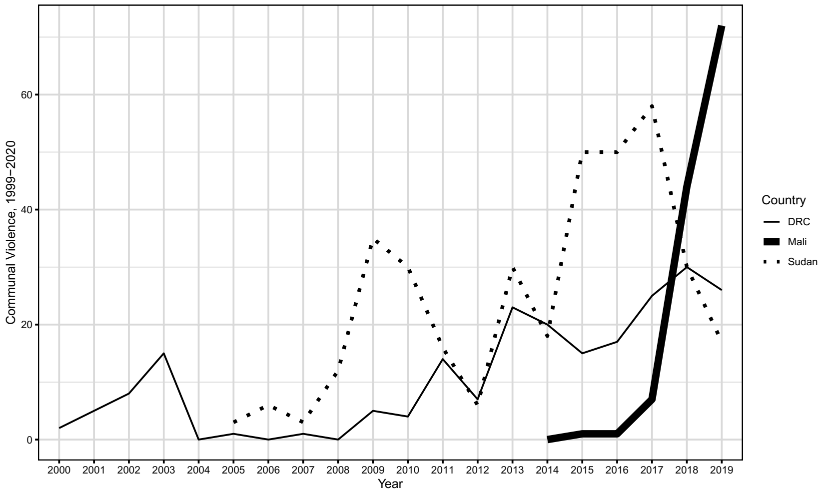


Figure 4.12 Variation over time in communal violence during UN peacekeeping deployments in Sudan, Mali, and the DRC

missions are more likely to deploy peacekeepers to particularly safe areas, then any negative correlation we observe between peacekeepers and the onset of violence may be spurious. *Ex ante*, I expect selection effects of this kind to bias our results downward toward zero because prior research has shown that peacekeepers generally deploy to the frontlines of a conflict (Fortna 2008; Ruggeri, Dorussen and Gizelis 2018).

However, to account for this type of potential bias, I model the temporal and spatial process that peacekeepers use to select which areas to patrol. Following convention, I use coarsened exact matching (Iacus, King and Porro 2012) to preprocess the sample along a set of factors that research and practice have shown determine where (and how) UN peacekeepers are deployed. I match grids that have peacekeepers deployed to those that do not but are similar in every other way. This procedure approximates an experimental setup in which peacekeepers are assigned as-if randomly to certain locations.

I measure travel time to major cities, distance to an international border, the grid's average mountainous elevation, population, predeployment levels of violence, nighttime luminosity, distance to the country's capital from the center of the grid, child mortality rate, predeployment levels of violence, and immunization rates.[20] These factors affect where

[20] Measures of accessibility are taken from AidData's GeoQuery tool. I use nighttime luminosity from the Defense Meteorological Program Operational Line-Scan System dataset of nighttime light emissions.

https://doi.org/10.1017/9781009432139.005 Published online by Cambridge University Press

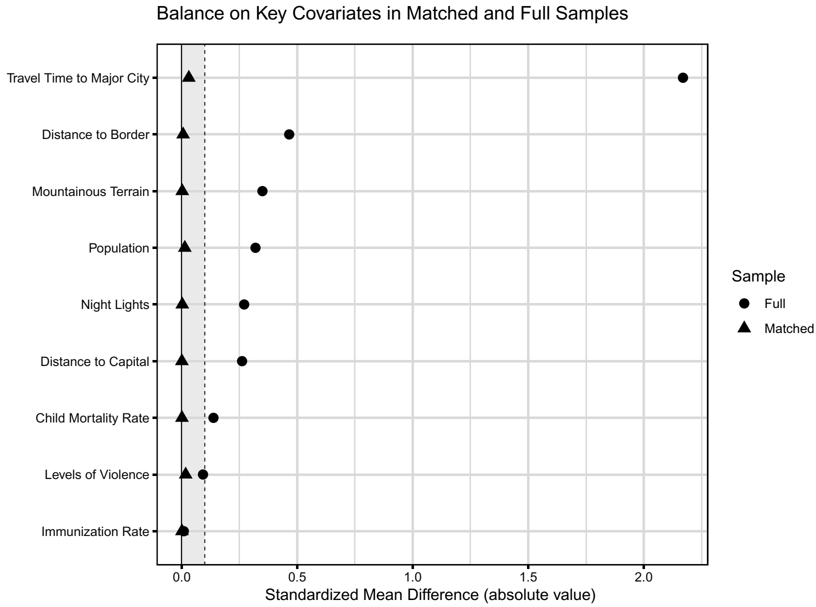


Figure 4.13 Standardized mean differences of key covariates between grids with vs. without peacekeepers (full and matched samples)

peacekeeping bases are initially sited because they predict both the need for peacekeeping to provide security and the logistical challenges the UN will face when establishing a base (Ruggeri, Dorussen and Gizelis 2018; Blair 2019). I also limit matches to within the same mission to ensure as close of a match as possible.

Figure 4.13 illustrates how much the matching procedure improves the sample by pruning it of observations in the data that would substantially bias it. The circles plot the standardized mean difference between grids that have peacekeepers and those that do not for the key covariates. Peacekeeping grids are different to a statistically significant extent (indicated by the shaded gray area) from nonpeacekeeping grids on all measures except immunization rates. Thus any comparison of grids in an unmatched sample would be akin to comparing apples and oranges. In the matched sample, there is no statistically significant difference between grids with and grids without peacekeepers. Although I cannot be certain that I have eliminated all potential sources of bias, this figure confirms that I have improved the balance of the sample comparison along several important observable dimensions.

In addition to preprocessing the data, I adjust the estimation for a set of time-variant and time-invariant factors that could be endogenous to the relationship between the local deployment of particular peacekeeping personnel and communal violence.

Communal violence could also systematically follow temporal patterns. For example, it may become more frequent as conflicts evolve because fragile states become progressively less able to resolve local disputes as state capacity disintegrates. For this reason, I include a cubic polynomial (t, t^2, and t^3) to explicitly model the temporal dependence of the units in the data (Carter and Signorino 2010).

I control for five variables that previous studies have associated with peacekeeping outcomes. First, I proxy for how violent a given administrative district is using a binary indicator of whether it had any civilian fatalities in the previous month. Second, since armed group involvement can quickly escalate a dispute, I proxy for the presence of armed groups using the number of battle-related fatalities in the grid in the previous month. Third, I include a dummy that indicates whether violence in a given district-month stems from an ongoing rebel conflict with the government. This helps me distinguish between violence generated by the escalation of communal disputes – my dependent variable of interest – and spillover violence from elite-level conflicts. Fourth, I control for the duration of a conflict since longer conflicts likely create new grievances between populations, making localized peace enforcement more challenging. Finally, I include a control for population since a larger population will have more opportunities for communal disputes.

Conclusion

The cross-national analysis permits me to estimate how deploying UN personnel affects the escalation of communal disputes across several different cases while controlling for factors that may otherwise be associated with violence against civilians. Moreover, given that I gathered the data from actual UN deployments and actual violent events, the data is closely tied to events on the ground as they unfolded. This gives me confidence in the external validity of these results.

However, observational data has at least three shortcomings, particularly across cases. First, the breadth of the data prevents me from analyzing individual deployments in depth, which makes it difficult to isolate any single causal mechanism. Second, the cross-national evidence presented is purely associational in nature; therefore it is difficult to entirely rule out threats to causal inference caused by endogeneity. Finally, operationalizing micro-level variables generates a significant loss of precision at the subnational level, which is exacerbated by including data from multiple cases.

To address these shortcomings, I begin with an in-depth examination of a single case, the UN peacekeeping mission in Mali. Focusing on a single case permits better data collection at a more fine-grained level. This allows me to distinguish between the existing explanations I highlighted in Chapter 3 and localized peace enforcement theory, which highlights the importance of perceptions of impartiality.

Part II begins with Chapter 5, which provides the historical detail necessary to understand the relevance of the Malian context. I explain that even within a single case, empirical analysis is difficult because peacekeepers seek out the most challenging disputes, locations, and conflicts. As a result, it is not straightforward to identify the causal effect of peacekeeping. In the empirical analysis that follows in Chapters 6 and 7, I use a multimethod approach that incorporates field experiments, surveys, a subnational statistical analysis of observational data, and interviews to offer a complete view of the case.