Security forces in conflict:

Composition, specialisation, and effectiveness in counterinsurgency and peacekeeping operations

Christoph Dworschak

A thesis submitted for the degree of

Doctor of Philosophy in Government

Department of Government

University of Essex

July 2020

Word count: 34,547

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am deeply grateful to all people who influenced the content of these pages, either directly or indirectly, and who supported me along the way. These are more people than I can possibly list here. To name a few however, I start with my supervisors at the University of Essex, Kristian S. Gleditsch and Federica Genovese. Kristian started to supervise me during my Master's, and Federica joined us in my second year. Their sharp foresight, rich perspective, and attentive counsel were essential to my academic development. At no point could I have wished for more thoughtful, committed, and supportive supervisors.

I owe special thanks to my supervising hosts Sabine C. Carey at the University of Mannheim and Jacob N. Shapiro at Princeton University for taking me in, and for their ingenious feedback and experienced guidance. For my third paper, I am first and foremost indebted to Deniz Cil for being such a splendid and reliable co-author, as well as to the peacekeepers who generously shared their experiences with us. I thank the colleagues and friends at the University of Essex, among which especially Tobias Böhmelt, Daina Chiba, Alejandro Quiroz Flores, and our PhD peer group contributed to my academic, didactic, and personal growth. Among the latter and though indebted to all, I wish to particularly name Marco N. Binetti for having been a brilliant office mate, reliable house mate, patient culinary instructor, and merciless chess player; and in the same breath I also wish to thank Metin Almisoglu, Melvin John, Marius Mehrl, Katariina Mustasilta, Marina Petrova, Christoph V. Steinert, and Julian Walterskirchen for all their suggestions, discussions, and pastoral care.

I am grateful for excellent feedback from many colleagues and friends at the Peace Research Institute Oslo, Princeton's Empirical Studies of Conflict project, the University of Konstanz, and the University of Mannheim. Special gratitude goes to Adam Scharpf for an impressive undergraduate course in which the very topic of this thesis, though unknowingly, first took shape, as well as for all discussions and exchanges that later followed. I also thank Belén González, Thomas Bräuninger, Moritz Marbach, and Stathis Kalyvas for their critical influence on my future path at the right moments.

I thank Jakob J. Gilg and Clara M. Hergert for many things that I cannot do justice to here, as well as all those friends who I did not get to mention already. Finally, I am grateful for my supportive family. As a first-generation academic, the importance of the roles that Felix, Gisela, and Wolfgang played along the way cannot be overstated. All my work and research is dedicated to the people affected by conflict. This manuscript is dedicated to my family who enabled me.

Abstract

How does the composition of security forces in counterinsurgency and peacekeeping operations influence their behaviour and performance? In this thesis I answer this question across three separate articles.

Modern state militaries are expected to perform numerous duties, including offensive combat in inter- and intra-state conflicts, policing, and state-building. In the first paper, I argue that a differentiation of the security apparatus allows for an increased specialisation in units' training and equipment on specific tasks, which enhances counterinsurgency effectiveness. However, fragmenting the security apparatus may also indicate coup-proofing efforts taken by the regime to mitigate military capabilities. I accommodate this ambivalence by devising a new index for differentiation, and find that differentiation increases counterinsurgency effectiveness.

Building on these findings, in the second article I focus specifically on security forces' defective behaviour, seeking to explain why security forces sometimes choose to side with the opposition movement. Drawing on the distinction established in my first paper, I argue that depending on the type of military fragmentation the risk of defection either increases or decreases. Employing Bayesian estimation, the findings in each paper provide corroboration.

The third paper provides an extension to peacekeeping operations. My co-author and I argue that peacekeeping troops greatly vary across two dimensions: their function specialisation and their country of origin. We argue that mixing different units greatly increase peacekeepers' specialisation in skills and equipment, improving peacekeeping effectiveness. However, this effect is strongly moderated by cultural diversity among troop contributing countries (TCCs), which exacerbates coordination problems. Our analysis on UN peacekeeping bases from 1994 to 2014 support our theory.

The three studies jointly contribute to the fundamental discussion on force composition, civil-military relations, and effectiveness in counterinsurgency and peacekeeping operations. Understanding these phenomena is pivotal to explaining how conflicts develop, escalate, and end.

1 Introduction

"*There is a lack of training, mostly of training, and then of course equipment too. Starts with ammunition, vehicles, and so on, but the basic thing should actually be education. There was [a] platoon, and they actually started their military service not even a year before they went on deployment, and three quarter of them got killed over night in an ambush. [...] When you establish a camp with your vehicles in the desert, you have to park the vehicles in a certain kind of way, and they didn't follow it and then it was quite easy to attack them and overrun them. Nobody wants to have a joint operation with them because that can be quite dangerous.*"

United Nations intelligence officer about their 2018 deployment.

This quote is taken from a conversation I had in October 2019, which was part of a series of interviews I conducted over the past months with peacekeepers of different backgrounds, ranks, and functions. United Nations (UN) peacekeeping offers a rare insight into the performance of troops from vastly different military organisations while being deployed to the same theatre of operation. Soldiers join UN missions with varying training routines, diverse civil-military traditions, and different perceptions of their own roles as "military professionals" (cf. Huntington 1957). Throughout all of my interviews it became evident that – unfortunately – it would be a mistake to treat the officer's story in the introductory quote as an exception. Instead, it must be understood as being representative of widespread, underlying patterns. These patterns are not only fundamental to UN peacekeeping, but underpin the inner workings of all military organisations. It is these patterns of security force composition and effectiveness that I seek to shed light on in this thesis.

There are four key lessons to learn from the officer's story: first, the officer emphasises the role of training and equipment. No matter the profession: having the right training and equipment usually constitutes a crucial prerequisite to fulfil one's task. The unique task of security forces is the exercise and management of violence (Huntington 1957). Second, because the peacekeepers lacked training in how to establish a camp, they themselves were overrun. The exercise of violence takes the shape of a "zero-sum game," which is to say that the degree by which one actor is inefficient in their exercise of violence,

e.g. through a lack in training and equipment, is equal to the advantage that their opponent will gain in their own exercise of violence. Therefore, when writing about security forces' effectiveness in pursuing their tasks, the stakes are very high and the "practical relevance" of this research topic can hardly be overstated. Third, the peacekeepers lacked oversight. If there was someone able to closely monitor their behaviour while out on patrol, the way they established their camp could have been influenced. Due to the nature of conflict, in all its severity and uncertainty ("fog of war"), monitoring the behaviour of troops is exceedingly difficult and constitutes a prime example of a so-called "agency problem" (Axelrod 1984; Akcinaroglu and Radziszewski 2013; Petersohn 2017). Fourth, the troops' ineffectiveness was not only to their own detriment, but their colleagues also felt endangered by their lack of training. The tasks of security forces are cooperative in nature, and any degree of ineffectiveness can have far reaching ripple effects on coordination. All of these points are core elements to this thesis.

This thesis is based on three separate articles. The first article, Dworschak (2019), theoretically establishes and empirically tests the intuition of the introductory quote: using security forces that lack the necessary equipment and training for their task at hand is not only inefficient, but detrimental to counterinsurgency effectiveness. I argue that having multiple units, each of which is specialised on a different security task, enables them to engage in division of labour. I call such a fragmentation of the security apparatus into multiple specialised units "differentiation." Asking "does force differentiation influence a state's counterinsurgency effectiveness?", I show that the more "differentiated" forces are into different functionally specialised units, the more security tasks can be covered by them. My second article, Dworschak (2020), provides an application of the first article's findings to civil-military relations: "How does military differentiation influence the likelihood of defections during (non-)violent uprisings?" Specifically, it argues that effectiveness contributes to the cohesion of security forces and minimises defections in the face of popular uprisings.

The third article is a co-authored project with Deniz Cil (University of Maryland), and is currently submitted to a journal for consideration. In lieu with the two preceding papers, the third one examines the effect of UN peacekeepers' composition on their effectiveness. Peacekeeping missions provide a particularly interesting setting to apply my concept of functional differentiation: troops greatly vary in skill and equipment, but operate in the same theatre. We argue that peacekeepers are more effective if they can draw om a diverse set of skills and equipment. This beneficial effect of functional diversity depends

however on seamless coordination. Therefore, when there is linguistic and cultural friction between peacekeepers, the beneficial effect of functional diversity disappears.

In the following I offer an overview of the relevant literature, distilling my theoretical concepts and my academic contribution overarching the topic of security force composition and effectiveness. I then review the sources of quantitative data which are employed throughout this thesis and discuss my methodology. After presenting the three papers, I conclude the thesis with a summary of my findings and illustrate various policy implications, ending with a discussion of future research avenues.

1.1 Effectiveness of security forces

Military effectiveness is a contested topic in the existing literature. Militaries perform many different tasks, like site protection, policing, offensive operations, and humanitarian assistance (Biddle 2004, 527). This task diversity makes it difficult to determine overall effectiveness. However, as mentioned above, militaries defining feature is the exercise and management of violence. Therefore, many scholars focus in their definition on the military's ability to inflict damage 'on the enemy's military while preserving [one's] own troops – holding situational [...] and environmental [...] factors, force size, and weapon technology constant' (Pilster and Böhmelt 2011, 333).

My first two articles build on this conceptualisation of military effectiveness, and extend it by considering the particularities of counterinsurgency warfare. Counterinsurgency operations require a carrot-and-stick approach, carefully combining the use of force with a "hearts and minds" approach on the operational level (Kalyvas 2006) and broader state building measures on the strategic level (Berman, Felter and Shapiro 2018). As such, my work is firmly embedded in the counterinsurgency research agenda which famously peaked with Kalyvas (2006) and more recently in the context of the US' counterinsurgency experience of the last two decades with Berman, Felter and Shapiro (2018), but which also reaches back to works like Galula (1964) and Welch (1976).

Counterinsurgency (COIN) operations differ from conventional warfare in one key feature: the enemy is not directly visible, and a state often has to wage war amidst and against its own population (Spear 2008, 400, 404). Consequently, the regime has to overcome an identification problem, determining who belongs to the enemy side (Böhmelt et al. 2020; Kalyvas 2006, 173). To overcome this identification problem, the counterinsurgent critically relies on popular support: in contrast to the

state security forces, there is little to none information asymmetry between the local population and the insurgents. Therefore denunciations by locals are paramount for the security forces to track down the location of rebel hide-outs, or to identify civilian collaborators (Berman, Felter and Shapiro 2018; Kalyvas 2006, 174-182; Kiras 2013, 179-180; Spear 2008, 397; Wickham-Crowley 1992, 7-8; Galula 1964, x). Such local cooperation with the security forces require winning over 'hearts and minds', i.e. inducing 'attitudinal shifts [...] through nonviolent persuasion [..., i.e.] programs of political liberalization, economic development, and civic action' (Kalyvas 2006, 94). For long-term stability after pacification was achieved, state building measures and political action are necessary to address the root causes of the insurgency (Berman, Felter and Shapiro 2018; Galula 1964, 88-89). However, as both noted by Berman, Felter and Shapiro (2018) and Kalyvas (2006), state building measures and winning hearts and minds also require a credible threat of force: if locals do not perceive the state security forces to be "in control" of their area and able to protect them from rebel retaliation, their willingness to disclose information is minimal. Therefore, to gain and hold territorial control, security forces need to both persuade civilians as well as deter opponents.

UN peacekeeping operations share important similarities with these counterinsurgency efforts. Today's peacekeepers face an increasingly diverse set of tasks, and these tasks became more "military" in their nature: mobile and site protection, separating and disarming combatants, capturing and holding checkpoints, patrolling, targeted interventions, and intelligence gathering. While in early missions peacekeepers were deployed mainly to monitor ceasefire treaties as mere observers, today's peacekeepers are deployed in countries with active conflicts and with mandates to pro-actively protect civilians through force engagement. These robust peacekeeping operations have the provision of security as part of their mandate, and usually involve conditions of power asymmetry requiring local intelligence (Howard and Dayal 2018; Hultman, J. Kathman and Shannon 2014, 2013; Fortna 2008). In other words, peacekeepers engage in conflicts in similar ways as counterinsurgents do (Pilster, Böhmelt and Tago 2016; Friis 2010; Celeski 2009; Lyall 2009; Kalyvas 2006), with the apparent exception of a purely offensive use of force (Howard 2019; Howard and Dayal 2018).

Similar to counterinsurgency operations, the variety in peacekeeping shows that local adaption, and a flexible set of skills and equipment are required. There is no particular toolbox of tactics and strategies that always has to be employed. Peacekeeping and counterinsurgency effectiveness rather

necessitate security forces' ability to select the right procedures, accounting for the specific characteristics of the security environment they face. To this end, security forces need to rely on local intelligence, and on appropriate training and equipment that are tailored to the security task they need to accomplish. Consequently, a unitary conventional force with a one-sided training background is unlikely to cover all bases (Bove and Ruggeri 2016; Pilster, Böhmelt and Tago 2016; Lyall 2009; Kalyvas 2006; Galula 1964, 66).

Importantly, security forces' effectiveness in handling intra-state conflict depends on their ability to adapt to the individual characteristics of the uprising they face, and their ability to make nuanced assessments about the use of force in each situation and theatre. Such an ability requires specialised training, equipment, and awareness of local cultures. This is difficult to achieve for a purely conventional, unitary force (Pilster, Böhmelt and Tago 2016; Lyall and Wilson 2009; Spear 2008, 390, 403-405; Welch 1976; Galula 1964, 66). Colonel Charles E. Callwell stated in 1906: 'Guerilla warfare is what *regular* armies always have most to dread, and when this is directed by a leader with a genius for war, an effective campaign becomes well-nigh impossible.' (Callwell 1996, 126; emphasis added).

1.2 Revisiting the composition of security forces

In the following I summarise existing scholarship on the role of force composition, and show how my thesis fundamentally contributes to these strands of literature. Based on the preceding discussion of effectiveness in counterinsurgency and peacekeeping operations, it appears intuitive that a composition in which a variety of is units available for deployment, all specialised in distinct security tasks, is preferable over one with only one unitary force. I develop a theoretical framework explaining how aspects of specialisation and division of labour between parallel units contribute to counterinsurgency and peacekeeping effectiveness. Beyond providing the first systematic investigation into these mechanisms, my framework also discusses important conditionalities of these relationships. In peacekeeping, coordination of functionally specialised units may be impeded due to linguistic and cultural diversity of UN troops. In counterinsurgencies, having multiple parallel forces may also be a sign of poor civil-military relations, which equally impedes coordination and hampers military capacity.

Starting with the role of force composition in counterinsurgency campaigns, there is a comprehensive body of literature suggesting that the way in which states structure their security forces is a

function of their civil-military relations (De Bruin 2020a; Dahl 2016b; Lutscher 2016; Roessler 2016; J. M. Powell 2012a; Svolik 2012; Belkin and Schofer 2005, 2003; Quinlivan 1999; Thompson 1976). Ailing civil-military relations may even peak in a coup d'état, i.e. in the military ousting the government. To minimise this risk, the political leadership can employ various coup-proofing techniques, one of which is called "counterbalancing:" based on the idea "divide and conquer" (Belkin and Schofer 2005, 147-148), the regime divides its security apparatus into multiple units with separate command structures, avoiding a concentration of power with only few commanders.

In addition, such a separation into parallel units raises coordination problems for coup-plotters: in case one military commander turns rogue and wants to topple the government, the leadership is able to deploy the remaining loyal parallel militaries against the renegade unit (to "counter-balance" it).[1] Such counterbalancing is generally accompanied by a set of other coup-proofing measures (Quinlivan 1999): to prevent coups, the leadership may hamper forces' capacity and resolve, and tends to foster military promotions not based on merit but instead on ethnic, religious, or family ties, in an attempt to ensure commanders' loyalties. Joint training exercises are also often limited, as they could help commanders in overcoming the purposefully induced coordination problems.

In summary, governments under high coup risk engage in coup-proofing techniques, which usually also involve counterbalancing.[2] With few exceptions, these coup-proofing techniques imply weakening the own security forces, either consciously or as a necessary externality. As Biddle and Long (2004, 532) show, less resolve, training, and capacity, as well as partisanship in the promotion of staff, lead to a decrease in human capital, military ineffectiveness, and generate frustration among soldiers and officers. Therefore it is unsurprising that counterbalancing is associated with a decrease in counterinsurgency effectiveness (Roessler 2016; J. M. Powell 2012a, 140-145). In addition, being a sign of poor civil-military relations and weakened resolve, counterbalancing also increases the risk of soldiers and officers defecting to the opposition movement (Dahl 2016a; Lutscher 2016).

This is where my first two papers come in, which systematically introduce the new concept of "differentiation." Contemporary literature missed that counterbalancing is not the only reason why security forces may be fragmented into parallel units. As described above, a government may also

1. See Belkin and Schofer (2003, 613), Belkin and Schofer (2005, 149-150), Pilster and Böhmelt (2011, 335-336), J. M. Powell (2012a, 40-41), and J. M. Powell (2012b, 1022-1023).

2. Sudduth (2017) argues that higher coup risk leads to less counterbalancing. However her coup risk concept is synonymous with the predicted probability of observing a coup, there coup-proofing is endogenous to it. I build on Belkin and Schofer (2003) in my conceptualisation of coup risk, who understand it as a latent variable capturing the underlying structural risk environment.

fragment its security apparatus into parallel units to allow for specialisation and a division of labour, thereby *increasing* counterinsurgency effectiveness. In other words, I argue that there are two types of security force fragmentation, counterbalancing and differentiation, which yield opposite implications for counterinsurgency effectiveness. While my first paper establishes this distinction and shows how the effect of security force fragmentation on counterinsurgency effectiveness is indeed heterogeneous, the second paper employs the same distinction to illustrate effect heterogeneity when it comes to security defections. These contributions substantially revise our understanding of the meaning and context of security force composition.

Turning to the role of force composition in UN peacekeeping operations (PKOs), recent scholarship shows that troops' country origins, and the way in which they are combined, have profound implications for PKO effectiveness (Bove, Ruffa and Ruggeri 2020; Bove and Ruggeri 2019; Haass and Ansorg 2018; Bove and Ruggeri 2016). Similar to security force composition in counterinsurgencies, Bove and Ruggeri (2016) argue that peacekeeping missions benefit from maintaining multiple parallel units from different troop contributing countries (TCCs) on a given base, as this enables peacekeepers to draw on a diverse set of skills and equipment.

On a day to day basis, peacekeeping operations require the coordination of various tasks performed by different unit types: engineers build the bases, repair equipment, and construct and maintain surrounding roads and bridges. Helicopter squadrons transport personnel and resources between bases, do air patrol, and support ground manoeuvres. Other important tasks are performed by medical support units, demining units, special forces, riverine units, reconnaissance, signals corps, etc. All of these unit types, called "enablers", serve to support infantry contingents in working to fulfil the mission mandate. Infantry contingents separate combatants and protect civilians through securing areas or sites, patrolling, dismantling checkpoints, and so forth. Following the discussion on effectiveness in the previous section, it is apparent that any lack of coordination between these different units severely impedes peacekeepers' force projection and reach.

My third paper, which is a collaborative project with Deniz Cil, is the first to systematically investigate the implications of combining these unit types. While previous research theorised about the importance of having different skills and equipment on a base, this was conceptualised and operationalised through the diversity in TCCs (Bove and Ruggeri 2016). While in reality the diversity in TCCs

is indeed positively correlated with the diversity in unit types, we show that it is useful to distinguish between these two concepts. Intuitively, and in line with my first two thesis papers, having different unit types on a base that are all specialised in different tasks substantively increases peacekeepers' effectiveness. However, when these units come from different countries, it becomes more difficult for them to coordinate their actions due to linguistic and cultural differences. Therefore, having multiple TCCs on a base is not a cause of increased effectiveness, but a mere placeholder proxying for the actual unit types present. This framework significantly contributes to recent scholarship on peacekeeping composition, while also illustrating the basic relevance of security force specialisation and effective coordination.

1.3 The empirical assessment of security forces

Capturing the composition of security forces is no easy endeavour. To date, there are only few data sources available. One of the most comprehensive ones for state security forces is the IISS dataset (International Institute for Strategic Studies 2016).[3] The use of these data have a strong tradition in the existing literature on civil-military relations (J. Powell 2019; Dahl 2016b; Lutscher 2016; Pilster, Böhmelt and Tago 2016; J. M. Powell 2012a; Pilster and Böhmelt 2012; Belkin and Schofer 2005). For this reason, and because it is the only source that records the size of each unit within a given security apparatus, I use them in my first and second paper for the design of my novel differentiation index. The unit of analysis is on the country-year level. In the past, studies on peacekeeping composition drew heavily on yearly country-level data by Hultman, J. Kathman and Shannon (2013). With the publication of Cil et al. (2020), base-specific monthly data on peacekeepers' numbers and TCCs became available. I also use information on peacekeeping unit types (infantry, engineers, etc.), which will be made publicly available in future versions of Cil et al. (2020).

For measuring my first paper's outcome variable, counterinsurgency effectiveness, I classify all counterinsurgency outcomes as win/draw/loss based on the UCDP Conflict Termination Dataset (Kreutz 2010). For my second paper's outcome variable, security defection, I draw on a binary measure from the NAVCO 2.0 data project (Chenoweth and Lewis 2013). In the third paper, following an understanding that peacekeepers require force projection and coordination in order to separate combatants (Cil 2019; own

3. With the others being data by De Bruin (2020b) and Carey, Colaresi and Mitchell (2016).

interview material), we measure peacekeeping effectiveness based on the occurrence of battle-related deaths.

The academic field of conflict studies has been comparably slow in adopting new methodologies and in fixing old empirical quirks. Trying to "be the change I want to see" in the field, I set myself high standards regarding my own empirical strategies. My first two papers make use of Bayesian rather than frequentist inference for multiple reasons. First, multilevel modelling is native to the Bayesian framework (McElreath 2015). Second, Bayesian inference estimates missing observations as parameters and adds them to the overall uncertainty (Little and Rubin 2002). Third, Bayesian inference allows for a sensible probabilistic interpretation of the results (McElreath 2015; P. M. Lee 2012; Gill 1999). In my third paper, I improve causal inference through the use of matching and simulations. The use of weights from Covariate Balancing Generalized Propensity Score (CBGPS) matching helps to reduce the correlation between observed confounding variables and the main treatment variable while relaxing functional form assumptions across strata (Fong, Hazlett and Imai 2018; Imai and Ratkovic 2014). The CBGPS matching approach is designed to overcome the dimensional limitations of traditional propensity score matching, and is suitable for continuous treatment variables. Meanwhile, simulations help to put bounds on the potential influence of unobserved confounders. Moreover, the third paper combines qualitative information with quantitative analyses, thereby drawing on the strengths of both approaches.

2 Divide and conquer:

Military differentiation and counterinsurgency effectiveness

Does force differentiation influence a state's counterinsurgency effectiveness? Modern state militaries are expected to perform numerous duties, including offensive combat, site protection, and state-building. I argue that a differentiation of the security apparatus allows for an increased specialisation in units' training and equipment on specific tasks, which enhances counterinsurgency effectiveness. However, shifting from a unitary army to a more segmented military may also indicate coup-proofing efforts taken by the regime to mitigate military capabilities. I accommodate this ambivalent influence of counterbalancing by devising a new index for differentiation. The results of a Bayesian ordered logistic regression provide empirical support.

Keywords: differentiation, civil-military relations, counterinsurgency effectiveness, coup-proofing

Article published at: https://ssrn.com/abstract=3523588

Replication data published at: https://doi.org/10.7910/DVN/KJLCZZ

2.1 Introduction

But if one will wholly apply himself to the making of Bows and Arrows,
whilst another provides Food, a third builds Huts, a fourth makes Garments,
and a fifth Utensils, they not only become useful to one another,
but the Callings and Employments themselves will [...] receive much greater Improvements,
than if all had been promiscuously follow'd by every one of the Five.
Cleo, in: The Fable of The Bees (Mandeville 1732)

This quote from Mandeville on the usefulness of the division of labour, even predating Adam Smith, is widely applicable (c.f. Smith 1976, Edwin Cannan's commentary). Although the implications of military sub-division have received little scholarly attention,[4] I will contend that such a "division of labour" within the state's security apparatus yields similar results as predicted by Mandeville and Smith: increased effectiveness. Pilster, Böhmelt and Tago (2016) argue that military differentiation may even serve as a direct and observable proxy for counterinsurgency effectiveness, but do not evaluate their claim. With this paper, I provide the first systematic theoretical and empirical investigation into the relationship between force differentiation and counterinsurgency effectiveness.

Intra-state conflicts constitute the overwhelming majority of violent disputes, and are the main cause of battle-related deaths around the globe (Melander, Pettersson and Themnér 2016). Traditionally, state militaries are trained and equipped to fight conventional inter-state wars. However, a comprehensive body of research suggests that insurgencies, due to their complex asymmetric nature, must be met with a nuanced combination of state-building, protection, combat, and policing. Considering this multitude of tasks that the security apparatus has to perform – a performance which is often further impeded by the regime constraining military training and equipment to mitigate the risk of coups d'état – a thorough examination of the implications of force structure for counterinsurgency campaigns is long overdue, and leads to the research question "does force differentiation influence a state's counterinsurgency effectiveness?".

I make three contributions to research on civil-military relations. Firstly, I link force differentiation and counterinsurgency effectiveness theoretically, while developing a rigorous conceptual distinction between differentiation and counterbalancing. These two types of fragmentation have never been sep-

4. With a few exceptions, including some earlier work; see e.g. Huntington (1957, 67).

arated in past research, even though they are likely to have contradictory influences on effectiveness. Secondly, I devise a novel measure for differentiation that is distinct from past counterbalancing indices, and that helps to set the two concepts apart empirically. Finally, my study adds to our general understanding of military strategy, and of success and failure in civil wars.

The paper is structured as follows. I first review existing literature on military effectiveness in intra-state conflict and military differentiation, identifying the research gap of how force differentiation influences a state's counterinsurgency effectiveness. Second, I offer an explanatory framework that connects military differentiation with effectiveness in counterinsurgency campaigns: differentiation of the military enables different units to specialise in distinct tasks, including asymmetric warfare. Adaptation of training, equipment, and strategy to the unconventional environment of insurgencies is crucial for a successful counter-campaign. These mechanisms are then illustrated using the Indian counterinsurgency approach in Punjab. From this theoretical framework I derive the hypothesis that a more differentiated security apparatus leads to a higher effectiveness in counterinsurgency campaigns. Third, I develop an empirical test based on a new measure for capturing differentiation. The results of a Bayesian ordered logistic regression indicate that differentiation indeed substantially increases counterinsurgency effectiveness. I conclude with a discussion of the findings and impact, as well as directions for future research.

2.2 Literature review

Does force differentiation influence a state's counterinsurgency effectiveness? To answer this research question, this section begins with a consideration of the different conceptualisations of military effectiveness in general, and counterinsurgency (COIN) effectiveness in particular. Secondly, military 'differentiation' is an uncommon concept for capturing a security apparatus' fragmentation, and must be demarcated from the better-known concept of 'counterbalancing'. This distinction is especially relevant because it is well-established that counterbalancing leads to *less* military effectiveness, while I will argue that differentiation, although strongly related to counterbalancing, bears the *opposite* implication.

2.2.1 Counterinsurgency effectiveness

How to best capture the concept of 'military effectiveness' is highly contested in the existing literature. Biddle (2004, 527) note that militaries are confronted with an abundance of different tasks, including

protection, policing, invading, and humanitarian assistance. The diversity of tasks makes it difficult to determine overall 'effectiveness'. Some propose a broader understanding, encompassing all war-related aspects including the strategic level (Talmadge 2015, 4). However, as each military's principal domain remains related to combat, many scholars converge on a more restrictive definition of military effectiveness, focusing on the ability to inflict damage 'on the enemy's military while preserving [one's] own troops – holding situational [...] and environmental [...] factors, force size, and weapon technology constant' (Pilster and Böhmelt 2011, 333).[5]

Military effectiveness in counterinsurgency operations is different from the conventional image of open warfare: the enemy is not directly visible, and a state has to wage a messy war amidst and against its own population (Spear 2008, 400, 404). Consequently, for successful COIN, the regime is critically reliant on popular support (Kiras 2013, 179-180; Spear 2008, 397; Wickham-Crowley 1992, 7-8). Effectiveness in counterinsurgencies is thus more complex than the simple definition of general military effectiveness outlined above, as 'inflicting damage on the opponent' requires the government to first *identify* who belongs to the enemy side (Kalyvas 2006, 173). To overcome this identification problem, gathering detailed native information through denunciations by locals is vital for the regime, in order to track down the location of rebel hide-outs, or civilian collaborators providing them with resources and equipment (174-182; Galula 1964, x). Such a degree of local cooperation with the regime critically depends on winning over the natives' 'hearts and minds', inducing 'attitudinal shifts [...] through nonviolent persuasion [..., i.e.] programs of political liberalization, economic development, and civic action' (Kalyvas 2006, 94; for the role of the political dimension, see also Galula (1964, 88-89), advocating a 4/1 ratio as in political/military). However, as Kalyvas (2006, 94-99) is careful to acknowledge, this attitudinal alignment is not enough to ensure actual cooperative *behaviour*, since existential needs and opportunistic desires have to be taken care of as well: a civilian who feels likely to be punished by the rebels will hardly disclose sensitive information to the government. Thus, a carrot-and-stick approach remains necessary to gain territorial control, persuading civilians and deterring opponents (cf. 94, footnote 8).

A regime that does not succeed in winning over hearts and minds and fails to solve the identification problem can no longer selectively target the insurgents and their support base. In many instances, this has led the military to resort to indiscriminate violence, like up-rooting entire villages to forcibly re-

5. For a more comprehensive discussion of the conceptual complexities, see e.g. Millett, Murray and Watman (1987), Brooks (2003), Biddle and Long (2004, 6), Biddle (2004, 527), Brooks (2006, 153), Pilster and Böhmelt (2011, 333-335), Talmadge (2015, 4-5), and Petersohn (2017, 3).

settle them, mass executions, or areal shelling (Kramer 2005; Lyall 2009; Zhukov 2015). As Downes (2007, 420) remarks, there is a consensus in the literature that any violence that does not distinguish between adversaries and civilians, i.e. does not alienate the general population from the risk of being targeted by the government, severely benefits the rebels' legitimacy, capabilities, and recruitment efforts (see e.g. Mason 2004, 150, 155-156; Kalyvas 2006, 151-152, 171-172; Lyall and Wilson 2009, 77-78; Pilster, Böhmelt and Tago 2016, 32-33). Still, Downes (2007, 420) finds that collective targeting turns out to be more effective for smaller scale insurgencies (geographically and support-wise).[6]

In summary, the military needs to make a nuanced assessment of the situation and choose its strategy and tactics carefully in order to achieve a swift and sustainable pacification. Therefore, counter-insurgency effectiveness does not refer to a particular toolbox of tactics and strategies that always ought to be employed, but rather to the military's ability to sensibly select the right procedures, accounting for the specific characteristics of the uprising it faces. Such an ability comes from specialised training, equipment, education in local cultures, mechanisms for swift adaptation, and comprehensive information collection, and can hardly be achieved by a purely conventional and hierarchical unitary force (Galula 1964, 66; Welch 1976; Spear 2008, 390, 403-405; Lyall and Wilson 2009; Pilster, Böhmelt and Tago 2016). As Colonel Charles E. Callwell stated as early as 1906: 'Guerilla warfare is what *regular* armies always have most to dread, and when this is directed by a leader with a genius for war, an effective campaign becomes well-nigh impossible.' (Callwell 1996, 126; emphasis added).

2.2.2 Differentiation versus counterbalancing

In order to illuminate the origins of my main explanatory concept 'military differentiation', I first have to take a step back and acknowledge that my outcome of interest, meeting the challenge of an insurgency, is part of the general objective of regime survival. Broadly speaking, there are two main violent threats to a ruler who wants to stay in power: one coming from 'the masses', i.e. the state's population or large groups within the society, and the other coming from 'the elite', which envelops high-ranking officials in the ruling apparatus and the military (Svolik 2012, 3-10).

I consider military differentiation as a partial explanation for successfully countering the first type of threat: insurgencies. As argued by Pilster, Böhmelt and Tago (2016), maintaining multiple parallel forces allows a government to specialise them in different areas. For example, and as outlined in the

6. See also Gentile (2013) for a critical review of 'hearts-and-minds', and Branch (2009) for a case study.

section above, COIN requires special training and equipment which is distinct from the skills necessary for waging conventional war. However, the causal pathway from differentiation to COIN success via specialisation has so far barely been acknowledged. Pilster, Böhmelt and Tago (2016) assume this relationship to explain the occurrence of genocides, but do not substantiate it theoretically or provide direct evidence. Practitioners have mentioned differentiation as a prerequisite for success in asymmetric warfare. For example, aforementioned Colonel Callwell (1996, 129-132) emphasises that enemies who engage in partisan warfare have to be met with equally mobile and small-scale units, dispersed across the area of operation, and definitely not with one large unitary army.

The existing literature commonly associates military fragmentation with the second type of threat: coups. Fragmentation may be used as a coup-proofing technique, known as 'counterbalancing'. The rationale behind counterbalancing is based on the notion 'divide and conquer' (cf. Belkin and Schofer 2005, 147-148), where the government separates its military into multiple distinct units with independent command structures in order to avoid concentrating control over the entire army in one place. This way, should a General choose to defect, the other unit(s) would stay functional and be able to counterbalance his attempt to overthrow the regime.[7] Importantly, Quinlivan (1999, 141-142) emphasises that the balancing security forces must rather be 'parallel militaries' instead of 'paramilitaries', referring to the necessity of them being 'large enough, loyal enough', and equally well equipped, in order to pose a credible threat to a deserting unit.

J. M. Powell (2012b, 12-14, 140-171) and J. Powell (2019) examine the implications of this coup-proofing technique for the onset of civil war. Several of the mechanisms invoke COIN effectiveness in order to explain why actors may be (dis-)incentivised to initiate a rebellion. Powell theorises that coup-proofing efforts of a regime render civil war onset more likely, as they would decrease the state's ability to meet the challenge. Regarding counterbalancing in particular, he concludes that its effect on COIN effectiveness is highly ambivalent, with both conducive and detrimental elements at work (J. M. Powell 2012b, 155-162). Especially under high coup risk, fragmented military units that are supposed to serve as counterweights are kept close to the capital, rendering civil war onset more likely, while under low coup risk the units' deployment potential decreases onset likelihood (J. Powell 2019). Pilster and Böhmelt (2011) also find an adverse effect of counterbalancing on inter-state military effectiveness.

7. For a more thorough explanation of counterbalancing, see e.g. Belkin and Schofer (2003, 613), Belkin and Schofer (2005, 149-150), Pilster and Böhmelt (2011, 335-336), and J. M. Powell (2012a, 1022-1023).

As J. M. Powell (2012b), they emphasise that under high coup risk (1) the direct command structure asserted for coup-proofing purposes denies commanders to flexibly adapt to the situation on the ground, (2) parallel militaries cannot engage in warfare as they are kept close to the capital in order to be able to deter potential coup plotters, (3) promotions in the military are based on loyalty instead of merit, and (4) joint training of the multiple forces is prohibited, impeding their deployability (Pilster and Böhmelt 2011, 335-337).

My main point of departure in this article is that governments may employ multiple parallel security forces, and thus fragment their security apparatus, *without* subjecting their militaries to any of the coup-proofing measures listed above. 'Counterbalancing' refers to fragmentation *for the purpose of coup-proofing*, implying both a fractionalisation of the security apparatus *and* coup-proofing measures. Therefore, while 'counterbalancing', where leaders attempt to render the military *less* capable, largely has a detrimental effect on COIN effectiveness, it does not follow that this extends to 'pure' military differentiation outside of coup-proofing efforts. I contend that governments also differentiate their military for *other* reasons than coup-proofing (e.g. for specialisation), and investigate how this influences COIN effectiveness.[8] In summary, I emphasise that the adverse effect of counterbalancing on effectiveness is not inherent to fragmentation of the security apparatus itself, but rather due to the concomitants of coup-proofing.

A major implication of this crucial distinction is that counterbalancing always assumes parallel forces to be 'large enough' to challenge a regular military organisation in case a coup is attempted (Quinlivan 1999, 141-142). For differentiation however, outside coup-proofing, there is no reason to assume the government would deliberately impede the militaries' effectiveness in any way, and the auxiliary forces do not need to be large. On the contrary, one might expect the size ratio between the regular and auxiliary militaries to be ever more unbalanced the more paramilitaries are *specialised*. This change in implications renders the use of a counterbalancing index for capturing 'differentiation' invalid – a fact that will be accommodated in the research design. It also sets this study's explanatory concept apart from J. Powell (2019), who convincingly argues that even units intended as counterweights, despite their common coup-proofing impediments, can serve as able counterinsurgents when there is low coup

8. An extensive account on the trade-off between forestalling coups versus insurgencies ('coup-civil war trap') is provided by Roessler (2016), and touched on by Svolik (2013). While my study adds to this discussion, an explicit comparative analysis of counterbalancing versus differentiation appears very promising for future projects. In addition, for reasons of scope this project is limited to a discussion of counterbalancing and specialisation, although there may be many more reasons for governments to fragment their military.

risk. The concept of differentiation offers an alternative mechanism linking fragmentation under low coup risk to counterinsurgency effectiveness, complementing the work of J. Powell (2019).

2.3 Theory

In this section I provide a theoretical link between differentiation and COIN effectiveness via specialisation, concluding with a short summary and the core hypothesis to be examined in the empirical analysis. To avoid ambiguities, I first define 'differentiation' and 'counterinsurgency effectiveness' more precisely.

Citing Boyle (2008, 183), 'insurgency' means the 'strategic use of violence by armed factions against a state or occupying force for the purpose of overthrowing the existing political order'. Due to the common power disparity between the rebels and the incumbent, as the latter controls the military and the police, conducting an insurgency normally implies fighting with asymmetric means (hit-and-run tactics, emphasising stealth and mobility; ibid.). I use the terms insurgency, rebellion, irregular or guerrilla war, and intra-state conflict interchangeably, since the more fine-grained distinction between these types of armed dispute captured by Boyle's definition are not relevant here (for others, see also Kiras 2013, 175-176; Joes 2011, 96-98; Spear 2008, 392). However, I refrain from the term 'civil war', as it may indicate a large-scale conflict in which all sides solely fight with regular and conventional means, which falls outside Boyle's concept.[9]

Counterinsurgency refers to the measures taken by the government or occupying force in order to overcome an armed rebellion. More specifically, counterinsurgency is defined as 'those military, paramilitary, political, economic, psychological, and civic actions taken by a government to defeat an insurgency' (392).[10] I understand a counterinsurgency campaign to be more "effective" if its propensity of leading to a termination of the uprising in favour of the government is higher.

Force differentiation, or military fractionalisation, is defined as the number and ratio of military organisations (cf. Belkin and Schofer 2003; Pilster and Böhmelt 2011; Pilster, Böhmelt and Tago 2016). These consist of the regular military (army), paramilitaries (also called 'auxiliary forces'; Böhmelt and Clayton 2018, 1), as well as relevant air and naval units.

9. This is highlighted by the high threshold of 1000 battle-related deaths commonly applied to measure a full-fledged civil war. See e.g. Collier et al. (2003), Collier and Hoeffler (2004, 565), and Fiala and Skaperdas (2011, 178).
10. Original definition by the US Department of Defense in 2007. For a discussion on how definitions differ across institutions, review Spear (2008, 392).

2.3.1 The fable of the bees: Specialisation

Per the introductory quote, workers' specialisation increases their task-specific experience and skill-level. I apply the same logic to military organisations, bearing theoretical and empirical implications for the effectiveness with which military tasks are performed. Countering an insurgency is one of these tasks, and although I expect the basic argument to hold in several other instances, I herein confine myself to the ramifications for COIN in particular, leaving any extensions to future projects.

In simplified terms, a state's security apparatus has to cope with external and internal threats: externally, it must demonstrate resolve to deter attacks or gain an advantage in negotiations, and if deterrence fails, it has to be able to coercively maintain the state's interests (e.g. defend its borders; see Fearon (1995) for terminology and theoretical grounding). Internally, security forces have to ensure the regime's survival by conducting counterinsurgency against popular uprisings, and by counterbalancing against elite defections (Svolik 2012, 3-10). These categories can be further partitioned into many different subtasks – as noted above, counterinsurgency is a multifarious venture that includes not only military action, but also policing and local state building (Pilster, Böhmelt and Tago 2016, 45; Celeski 2009; Kalyvas 2006, 94-99; Brovkin 1994, 91).

A single security organisation may be used to execute all of these tasks. However, a regular army equipped and trained for conventional warfare might prove more harmful than helpful in the context of asymmetric conflict: it is inapt to selectively target enemies, or to co-opt the local population (Pilster, Böhmelt and Tago 2016, 31-39; Goswami 2015, 168-171). Similarly, an auxiliary force that is only lightly armed and intended for civil pacification would probably not perform well against another state's mechanised army. Consequently, a regime must choose how to train and equip its security organisation, and strike a balance between conventional and asymmetric specialisation, with decreasing returns towards both extremes. Dividing it into multiple units however, with at least one for each specialised task, can avoid this sharp trade-off. The usefulness of such task-sharing is further amplified since many security issues are heavily intertwined, and often faced simultaneously (cf. Gleditsch, Salehyan and Schultz 2008).

Another way to conceptualise this approach draws on general risk assessment: a government must assess the manifold security risks it faces, and design its security framework accordingly.[11] Both

11. See e.g. Huntington (1957, 66-68). 'Risk' is commonly understood as an event's probability to occur, times the magnitude of its consequences (Larson and Peters 2001, 26-27; Heng 2003, 79).

internal and external challenges have a certain likelihood,and both can be fatal for an unprepared government (King and Zeng 2001, 7). Therefore, the utility of employing the first military organisation in one field of specialisation is greatest, and the government's marginal utility of each additional force within the same field experiences diminishing returns. Thus, assuming the incumbent to be a utility-maximising actor, increased military differentiation indicates a higher probability of encountering specialisation.

Lastly, shifting the perspective from the government to the security forces as agents, fractionalisation benefits effectiveness by fostering military adaptiveness: with a growing number of parallel forces, the incumbent's ability to micro-manage all of them declines. As recognised by Avant (2007, 83-84) for British forces, looser political control gives military units greater leverage over their choices of strategy and tactics, which enables them to flexibly accommodate the situation on the ground, increasing their effectiveness.[12] Pilster, Böhmelt and Tago (2016, 45) also hint at the relevance of the security forces being able to learn from and adapt to local challenges in order to foster effectiveness. A real-life example for this mechanism constitutes the so-called 'wider peacekeeping' approach adopted by the British army in Bosnia, which entailed more independence, flattened hierarchies, and less doctrine, and eventually led to increased effectiveness (Thornton 2000; see also Spear 2008, 400).[13]

2.3.2 Illustration: The case of Punjab

In the region of Punjab, the Indian government used an exemplary variety of (para-)military forces and asymmetric strategies to counter the Sikh insurgency from 1984-1994 (cf. Gates and Roy 2016; Ray 2007; G. Singh 2000, 162-174). Below, I highlight how the differentiation of the security apparatus allowed for specialisation, which in turn enabled complementary task sharing. The narrative is not meant as an empirical test, but helps to illustrate my mechanism. Also, it entails a useful counterexample of what happens when a merely conventional strategy is adopted for a COIN operation, oblivious to local culture and political cleavages.

The decade-long violent unrest started in June 1984: 'Operation Blue Star' constituted a major military crackdown on a Sikh militant group led by Jarnail Singh Bhindranwale, based in the Golden Temple complex in Amritsar, Punjab (127-130). The operation was ordered by Indian Prime Minister

12. In her study however, the political control is loosened by the principal not being a unitary actor, and not by an environment of multiple agents. The fact that her findings fit my propositions nicely underlines how this in-depth examination of differentiation amends existing scholarly work.

13. This operational flexibility shall not be confused with 'mission tactics', which are also highly influential when it comes to effectiveness, but do not necessarily correlate with looser political control (cf. Van Creveld 2005).

Indira Gandhi and conducted by the Indian army,[14] starting with a curfew, military occupation, and total shutdown of the entire Punjab region lasting for several days (Ahmed 1996, 130; Brar 2000, 54). During this time, the security forces mounted an attack on the Golden Temple complex using heavy weaponry (including mortars), seemingly ignorant to the fact that it is an important centre of Sikh religion and culture, as well as on houses of Sikh worship throughout Punjab to detain other dissidents (Silva, Marwaha and Klingner 2009, 10; Telford 2001). While it resulted in the killing of Bhindranwale and his militants, it also caused the deaths of several hundred civilians 'caught in the fire' (Pillalamarri 2014; Telford 2001; G. Singh 2000, 163-165). This blunt approach and desecration sparked outrage among the Sikh community and widespread mutinies from Sikh members of the security forces, and eventually led to the assassination of Prime Minister Indira Gandhi by her Sikh body guards (S. Singh 2002, 382-384; G. Singh 2000, 129). The assassination in turn led to the retaliatory, infamous 'anti-Sikh riots' resulting in thousands of casualties among the Sikhs (Pillalamarri 2014).

A decade-long struggle against a network of Sikh militant groups followed, which eventually was overcome by a combination of political measures and a well-designed COIN strategy, involving a variety of different security forces (Silva, Marwaha and Klingner 2009, 11; G. Singh 2000, 162-174). To conduct this efficient counterinsurgency operation, and to avoid repeating the mistakes of Operation Blue Star, the Indian government changed its approach significantly: command was shifted from the federal army to the regional security forces under police chief Kanwar Pal Singh Gill (165-166). Especially for gathering local intelligence on rebel networks, regionally embedded security forces are essential (Pilster, Böhmelt and Tago 2016, 32; Goswami 2015, 168-169). Accordingly, Punjabi armed forces infiltrated the militant groups on all levels with special hit squads, exploiting local structures and social hierarchies, and brought them down with a mixture of targeted killings and selective violence (G. Singh 2000, 134). Gill's approach became known for its efficiency, cunningness, and subtlety, but also for human rights violations like torture and extra-judicial killings (Silva, Marwaha and Klingner 2009, 11-12). This was possible since the security personnel were given 'free hand' in their strategic and operational choices, enabling them to flexibly adapt to current circumstances (G. Singh 2000, 170-172).[15]

14. Though paramilitary units took part in it, they were placed under direct command of the regular military.

15. While this undermining of state legitimacy and democratic values may also impede COIN effectiveness, such a detrimental effect seems to have been outweighed by the benefits of circumstantial flexibility in this particular case. However, an in-depth discussion of how these contrasting influences played out goes beyond the scope of this paper.

In addition, the Indian government deployed paramilitaries, special forces, and federal army units to the region in support of the local units, which nevertheless remained firmly in charge (G. Singh 2000, 165-167). This boosted their firepower and served as a counterweight to reports of corruption: the downside of local embeddedness is that individual units may be prone to shirking and collusion (Pilster, Böhmelt and Tago 2016, 31). Consequently, the COIN operation heavily benefited from the simultaneous deployment of forces with different degrees of local integration: while the Punjabi security forces were most efficient in the elimination of insurgents and militant leaders, they lost the trust of the civilian population due to allegations of corruption. The presence of (impartial) federal forces proved helpful, since locals turned to them for protection and collaboration (G. Singh 2000, 167). This complementary task assignment was only possible because many of the federal (para-)military units were also specially trained in policing and counterinsurgency, having been previously deployed to the Assam region (167).

In summary, India's initial approach in fighting the militants was based on conventional means, was insensitive to culture, and showed no attempts to co-opt the civilian population. However, having learned from this botched operation, the subsequent insurgency was overcome much more efficiently. First, command was shifted to the regional level to increase available intelligence, and enable selective targeting. Second, a diverse set of security forces was deployed to accommodate different tasks: from mingling with the militants for gathering information, to killing insurgent leaders, to cooperation with civilians.

India's security apparatus consists of many parallel forces. Conventional wisdom dictates that such high fragmentation is motivated by coup-proofing: military fractionalisation is commonly understood as an attempt to counterbalance and restrain military forces in light of a salient coup risk. However, India is one of many cases with high fragmentation despite a *low* coup risk,[16] and the example above illustrates how it may serve other purposes, including specialisation, task sharing, and operational flexibility. These traits bear fundamentally different implications for a country's military effectiveness, and allow waging more efficient counterinsurgency campaigns. Thus, next to India's inherent need for (signalling) conventional capacity due to its ongoing inter-state rivalry and risk of conflict with Pakistan, it also maintains considerable, parallel asymmetric capabilities.

16. Its mean coup risk index value is -1.04. For scale, see the paper's summary statistics.

2.3.3 Hypothesis

In conclusion, differentiation leads to task sharing within the security apparatus, with units specialising in distinct strategies, equipment, and training, increasing the state's capacity to effectively counter an insurgency. It also fosters combat flexibility and strategic adaptation as operational agency is increasingly shifted to the security forces. This yields the hypothesis below. Note that, in line with my theory and conceptualisation, differentiation assumes low coup risk. Therefore, by using the term "differentiation", the hypothesis implies an interaction.

H_1*: Greater differentiation of a state's security apparatus leads to a higher effectiveness in counterinsurgency campaigns.*

2.4 Research design and empirical evidence

This study focuses on the macro-level relationship between the two main concepts, and the marginal effect of military differentiation on a state's ability to ensure its internal pacification. I first discuss how to operationalise differentiation and COIN effectiveness, and review the available data sources. I then present the method for the quantitative analysis, and examine and discuss the analysis' results.

2.4.1 Operationalisation and method

For the dependent variable, classify COIN outcomes as win/draw/loss. These classifications are obtainable for all intra-state conflicts through the UCDP Conflict Termination Dataset (Kreutz 2010).[17] Several issues arise from this measure: firstly, the measure's reliability may suffer from coding subjectivity. However, since potential coding errors are not systematically related to differentiation, they are unlikely to skew the results, even if their efficiency is slightly impaired. Secondly, the face validity COIN outcomes as a measure for COIN effectiveness could be challenged because such a final outcome is shaped by many other situational and environmental factors other than military effectiveness. Hence, although it is a more 'general' measure than e.g. a loss-exchange ratio, and less sensitive to resource-intense tactics,

17. Where one side is a government whilst the other is not, and based on UCDP inclusion criteria (threshold of 25 battle-related deaths, among others – see the codebook (Kreutz 2010) for details. I deem this superior to using a loss-exchange ratio (LER; cf. Biddle 2004, 535; Pilster and Böhmelt 2011, 338), because there are potentially effective COIN strategies that deliberately accept high losses. For example, a government may choose to put soldiers at risk by deploying large numbers of infantry in order to enhance grass-root intelligence (Lyall and Wilson 2009). Most importantly however, such data are not globally available for intra-state conflicts: at the time of this writing, the battles for which data on dyadic casualties were collected only pertain to inter-state wars (Biddle 2004; Pilster and Böhmelt 2011).

it can also be more susceptible to confounding influences. In order to accommodate these factors, I introduce several control variables that may be associated with both outcome and differentiation below.

Lastly, the distribution of UCDP outcome levels is highly imbalanced, and what constitutes a conflict's 'end' (or whether a government was militarily dominant) is ambiguous. For example, only a small portion of UCDP outcome levels are actual decisive 'victories' for either government or rebels. Other outcomes that dominate the sample are ceasefires, peace agreements, and low activity (i.e. the dispute remained unresolved, but battle-related deaths fell below the threshold of 25). This imbalance is in line with the argument that insurgencies are a form of 'new wars' commonly associated with the absence of decisive victories (Williams 2008, 165-166). Both 'low activity' and 'ceasefire' are outcomes indicating that the dispute remained unresolved. Following Cunningham, Gleditsch and Salehyan (2009, 583) and Wood and Kathman (2014, 695), I code such unresolved conflicts that have recurred within three years between the same actors as a single conflict.[18] Those remaining conflicts that do not recur, but actually 'end' in low activity, are dropped from the sample.[19]

Based on the preceding definition 'ground-combat compatible military organisations', Pilster, Böhmelt and Tago (2016) operationalise differentiation with an index that takes the security units' quantity and respective sizes into account:

$$\frac{1}{\Sigma_j s_{jit}^2} \tag{1}$$

with j being the number of separate military units per country-year observation (i and t), and s being each separate unit's share of overall military personnel within an observation. A higher index value thus indicates higher differentiation, with 1 being the baseline (no differentiation). A 'unit' may refer to any entity of the state security apparatus with an independent command structure to the executive, including the regular army, militias, presidential and border guards, militarised police, and special forces.

This index was initially conceived by Pilster and Böhmelt (2011) to capture the concept of counterbalancing. Following the conceptual distinction I emphasise on page 23, using a counterbalancing index for measuring differentiation bears two major problems. First, counterbalancing requires units to

18. The time-span within which recurrence is coded as a continuation of the same conflict varies in the literature. Changes along these lines do not impact my findings' robustness.

19. Following Howard and Stark (2017, 142-143), I also considered such a termination as a form of 'stabilisation' in favour of the government, coding them as government success. In these cases, quoting Sullivan, Patricia L. and Karreth (2015, 275-276), rebels 'are degraded to the point that they can no longer conduct significant attacks.' This conceptualisation however is not uncontested; for a discussion, see e.g. Staniland (2017, 465) who argues that '[l]ow numbers of dead bodies do not equate to [...] centralized control over violence [...].' Either way, this change to the data did not impede the findings' robustness.

Table 1: Data summary

	Data Type	Descriptive Statistics
Dependent Variables		
Outcome	ordinal	1 = Rebel success (n: 7), 2 = Draw (n: 42) 3 = Government success (n: 24)
Explanatory Variables		
Diff. Index	ratio	Range: 3.00 (none) – 7.82 (max) Mean: 4.25, Median: 4.02
Coup Risk	ratio	Range: -4.47 (min) – 3.57 (max) Mean: 0.29, Median: -0.02
Other Variables		
GDP/c	ratio, log	Range: 6.44 (min) – 10.15 (max) Mean: 7.87, Median: 7.68
Polity2	ordinal	Range: -9 (auth.) – 10 (dem.) Mean: 0.32, Median: -1
Mil. Exp.	ratio, log	Range: 0.09 (min) – 2.81 (max) Mean: 1.25, Median: 1.20
Oil/c	ratio, log	Range: 0.00 (min) – 1.98 (max) Mean: 0.30, Median: 0.01

be large enough to pose a credible threat to potential coup-plotters.[20] This idea is also reflected by the index: imagine a country with two military organisations in a given year. If they are of equal size, the index shows a value of $1/(0.5^2 + 0.5^2) = 2$. Meanwhile, if one is a large conventional force of 99000 soldiers, and the other a specialised paramilitary of 1000, it is $1/(0.99^2 + 0.01^2) = 1.02$ (hardly any differentiation). It is apparent that size disparity is severely penalised by the original index' properties, which renders it useful as a measure for counterbalancing, but highly undesirable when trying to draw a link between differentiation and *specialisation*, independent of coup-proofing (see again page 23). Therefore, I adjust the index to instead reward unequal sizes:[21]

$$2\sqrt{j} + \Sigma_j s_{jit}^2 \tag{2}$$

20. With new data by De Bruin (2018) that captures unit functions, this assumption can be relaxed in future research.

21. The square root attenuates the increase in the index across units for higher values of j (diminishing returns), as it seems intuitive to assume that an increase from having one unit to having two units is more impactful than an increase from 10 to 11 units. See Online Appendix A for a visualisation. The estimation results are robust however to many different versions of this index, including simple linear versions. The necessary raw data for the re-calculation was kindly provided by Tobias Böhmelt, and originally retrieved from the International Institute for Strategic Studies (2016) and Clodfelter (2008). IISS data is known for inaccuracies and inconsistencies, and recent scholarship started to re-examine coup-proofing concepts using new data (De Bruin 2018). Therefore, while this project started using IISS data in 2015 in order to inform the majority of counterbalancing literature, future replications of these findings using other data and operationalisations are worthwhile.

The second problem originates from the fact that fragmentation for the ends of coup-proofing (counterbalancing) is associated with a substantial decrease in military capabilities and deployability (as discussed on pages 23 and 23). This leads me to introduce my first control variable: structural coup risk. Following Carey, Colaresi and Mitchell (2016, 64-68), an increased risk to experience a coup d'état leads a government to counterbalance, and to bolster its other coup-proofing measures. Therefore, I expect differentiation under high coup risk (i.e., counterbalancing) to be detrimental for COIN effectiveness, compared to differentiation under low coup risk. The variable Coup Risk is taken from Carey, Colaresi and Mitchell (2016), consisting of indicators for the strength of civil society, state legitimacy, and past coups. For a discussion of these components, and the advantages of capturing structural coup risk instead of mere propensity, refer to Belkin and Schofer (2003). I interact it with the adjusted differentiation index due to the expected change in the latter's slope over the range of realisations of Coup Risk – in other words, the interaction term between Coup Risk and the Differentiation Index constitutes the analyses' centre of interest. Recent research by J. Powell (2019) employs an interaction between counterbalancing and coup risk for similar reasons. I contend that due to the severely contradicting implications of differentiation versus counterbalancing, these two amendments are highly advisable when studying either of them, enabling the researcher to tell them apart empirically.

Other control variables are GDP per capita (logged) and military expenditure, both of which serve as indicators for state capacity and potential force projection, which may easily translate into the state's ability to differentiate its military and its vulnerability to insurgencies. The Polity2 index, classifying each country in a given year on a scale from total autocracy to full democracy, and a state's oil revenues per capita (logged), are included to control for regime legitimacy and reliance on popular support, both oh which are expected to systematically vary with the ability to wage asymmetric war and security force structure. All these variables are taken from the replication data for Carey, Colaresi and Mitchell (2016).

The final data contains 73 observations, from 43 countries between 1981 and 2000. The unit of analysis is individual intra-state conflict dyads, in the year in which they have experienced their respective outcome (right-censored cases are omitted for their lack of an outcome level). All variables are measured at the final year of conflict. Since the dependent variable is ordinal, I conduct an ordered regression using

logit links, assuming parallel predictors across categories. This corresponds to a simple proportional odds model (Ari and Zeki 2014; Agresti 2002).

$$P(Y \leq y_j^* | \boldsymbol{x}) = \left[\frac{exp(a_j - \boldsymbol{x'\beta})}{1 + exp(a_j - \boldsymbol{x'\beta})} \right],$$

with a_j as intercepts and $\boldsymbol{x}$ being a matrix of the independent variables Differentiation, Coup Risk, (Differentiation * Coup Risk), Gdp/c, Polity2, Military Expenditure, and Oil/c. I choose to run a Bayesian model due to the ease of probabilistic interpretation (McElreath 2015, 4-17; P. M. Lee 2012, 139-140; Gill 1999) and to avoid the problematic accept/reject logic of frequentist hypothesis testing (Gill 2018). One more particular advantage is the fact that Bayesian estimation does not rely on asymptotics, rendering it more feasible for small sample sizes (McNeish 2016). For estimation, I use JAGS and non-informative (flat) priors.

2.4.2 Results

Table 2: JAGS ordered logistic regression

	Post. means	Post. SD	Percent positive
Differentiation Index	-0.458	0.345	9.13%
Coup Risk	1.471	1.011	92.68%
Diff.$*$Coup Risk	-0.391	0.254	5.85%
GDP/c	-0.577	0.494	11.2%
Polity2	-0.026	0.055	32.58%
Military Exp.	0.536	0.563	82.85%
Oil/c	1.325	0.668	98.20%
j_1: Rebel win | Draw	-7.952	4.374	2.72%
j_2: Draw | Gov. win	-4.493	4.281	13.70%

N chains: 3, Burn-in: 5000, Thinning: 20, Eff. sample: 2000
Mean deviance: 134.08
N obs: 73

"Post. means" denotes the means of the posterior distributions. In a frequentist sense, these are the coefficients of the ordered logistic regression. "Post. SD" are the standard deviations of the posterior distributions. "Percent positive" facilitates the interpretation of the estimates' (un)certainty by showing their probability of being larger than zero.

Table 2 on the preceding page shows the results of the ordered logistic regression. The direction of the parameter estimates of interest are in line with my theoretical expectations, and experience low uncertainty despite the small sample size. The probabilities of differentiation and the interaction term to be non-negative are 9.13% and 5.85%. Similarly low, the probability that coup risk is negative is 7.32%. For a visual interpretation of the "coefficients", see Figure 1 on the next page.

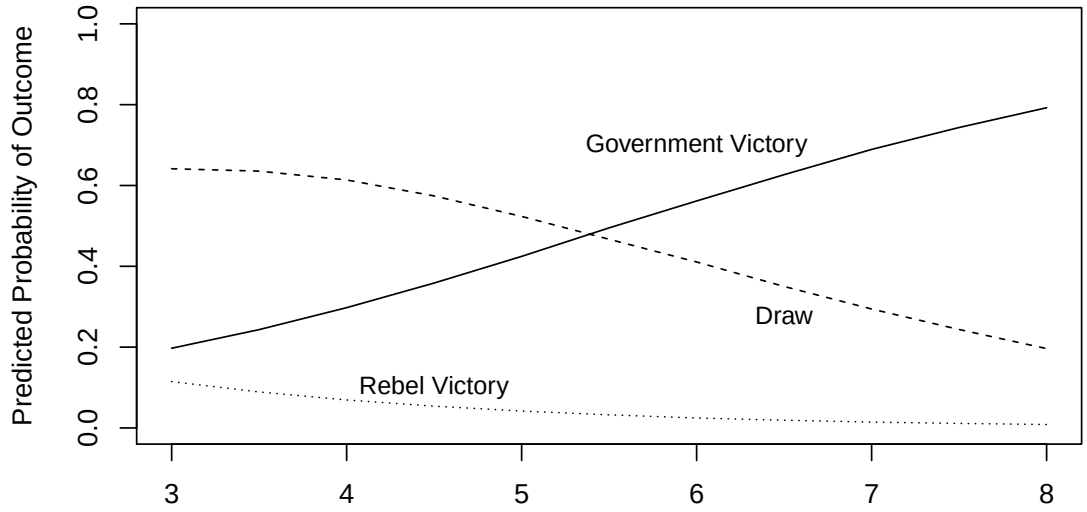
The figure renders apparent that differentiation exerts a sizeable effect on the predicted probability of observing either conflict outcome.[22] Comparing a setting with no differentiation to a setting with maximum differentiation, the predicted probability of government victory increases from 20% to 80%. Put differently, fragmentation that is not caused by coup-proofing (then: differentiation) makes a successful COIN campaign of the state's security apparatus more likely. For reference, the lower half of the graph shows the relationship under high coup risk, which indicates the expected opposite effect. In other words, whenever a regime is most inclined to make coup-proofing efforts, increased fragmentation (then: counterbalancing) seems to render the successful conduct of COIN campaigns more difficult. While my index is adjusted to capture differentiation rather than counterbalancing, and while I investigate conflict outcomes rather than onset, these results are in line with and lend further corroboration to the recent findings by J. Powell (2019).

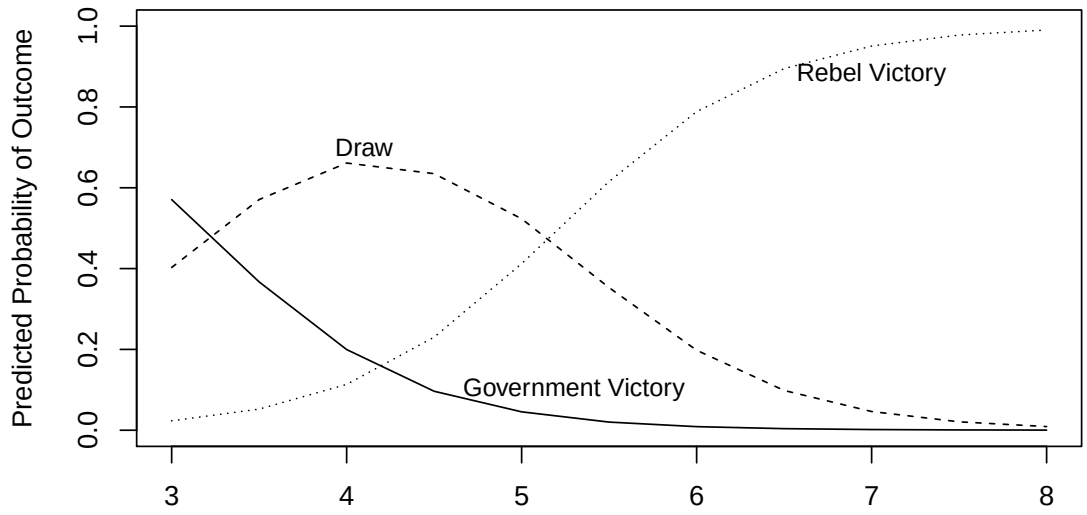
Credible intervals are omitted to increase the graph's readability – instead, the results' uncertainty at minimum and maximum values of coup risk is visualised separately, in Figure 2 on page 36. It shows the coefficient distributions of the differentiation index for minimum and maximum values of coup risk. It is apparent that the marginal effect of differentiation, that is the differentiation index under low coup risk, is positive with a 92% probability (8% is at or below zero). The joint overlap between the marginal effect of differentiation under high and low coup risk is 7.9% of the two areas, indicating that there is a substantively meaningful difference.

In summary, the empirical assessment supports the logic I have put forward in the literature review and theory section. More specifically, the results suggest that the hypothesised relationship, i.e. an increased differentiation of a state's security apparatus leads to a higher effectiveness in counterinsurgency campaigns, is indeed reflected in the sample. Furthermore, the conceptual distinction

22. Coup risk is held at its 5th (low, -2.6) and 95th (high, 3.2) percentiles. All other controls are held constant at their medians.

Figure 1: Predicted probability of outcome

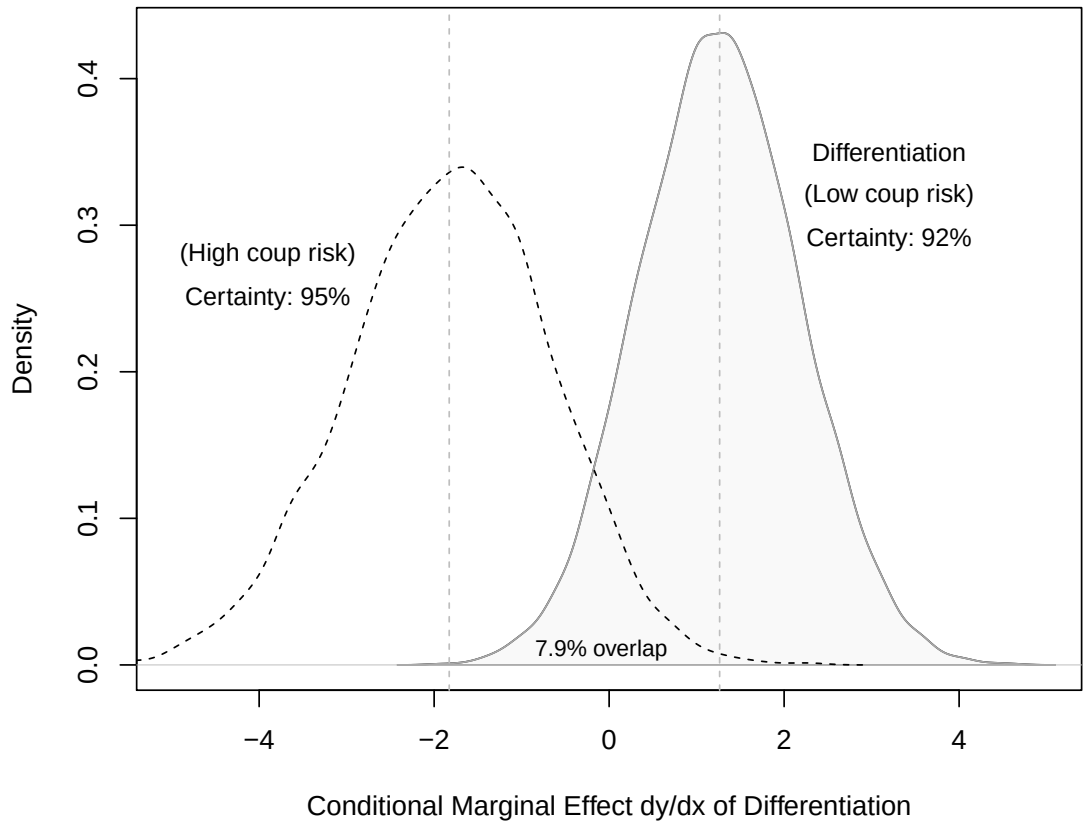




and proposed ambivalence between differentiation and counterbalancing is empirically signified by the sizeable mediating effect of coup risk.[23]

23. Online Appendix B offers additional results with more control variables, a squared term of Polity2, and a *null* model. Inference remains largely unimpaired.

Figure 2: Posterior coefficient distributions



2.5 Conclusion

I have examined whether force differentiation influences a state's counterinsurgency effectiveness, and identified force specialisation as a causal pathway. Differentiation of the security forces increases the likelihood of specialisation, and enables governments to assign individual units to pursue distinct tasks. This in turn facilitates appropriate training and equipping of the military apparatus for accommodating both external and internal threats, and thus fosters its capacity to effectively counter insurgencies. Additionally, differentiation increases the chances of commanders having more autonomy during an operation, rendering their choice of strategies and tactics more independent and flexible. I illustrated this mechanism using the Punjabi counterinsurgency operation as an example, and evaluated my hypothesis that differentiation leads to increased counterinsurgency effectiveness by employing a new measure of

differentiation independent of coup-proofing effects. A Bayesian ordered logistic analysis of differentiation on conflict outcome yields results that are consistent with my claim.

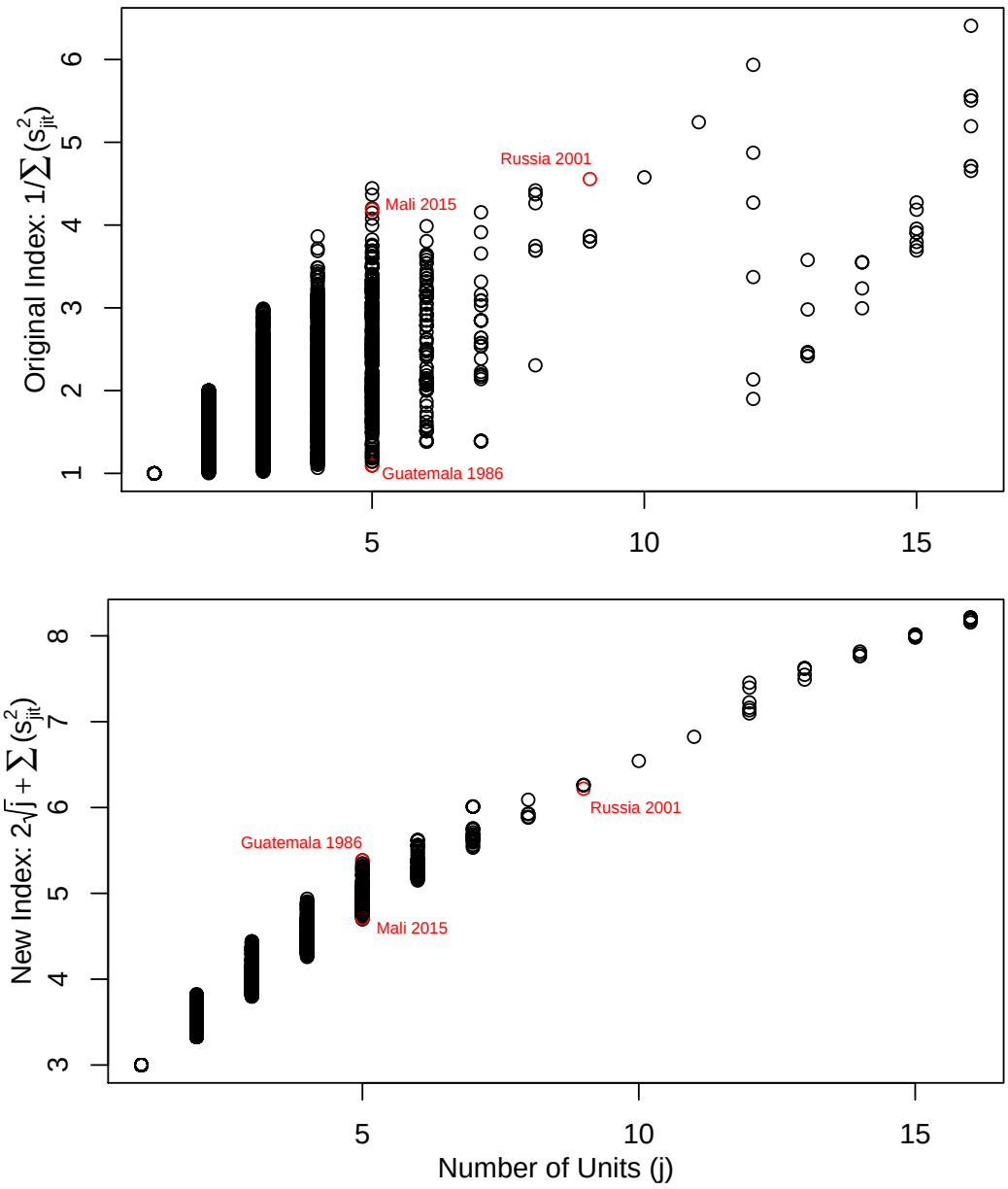
There are some weaknesses to this study that future research may remedy: First, the operationalisation of COIN effectiveness is an issue without a single 'best' solution, and subsequent work can consider alternative measures. Second, there is an empirical implication of my theoretical model that would be interesting to examine, and that may additionally substantiate or undermine my framework: counterbalancing implies, other than mere differentiation without coup risk, that the parallel forces are held near the nation's capital, and that the leadership is reluctant to deploy them to fight remote insurgencies (J. Powell 2019; J. M. Powell 2012b). If differentiation and counterbalancing indeed have different implications for COIN effectiveness, fragmentation under high coup risk would connote that military effectiveness in and around the capital may be higher. While testing this is beyond the scope of this paper, the idea appears very promising for future studies, and has become feasible with the increasing availability of geo-coded data (e.g. ACLED; Raleigh et al. 2010). Third, the latent nature of the core concepts and limited data availability call for a designated qualitative approach with greater depth to enhance internal validity. For example, looking at UN peacekeeping missions provides interesting insights into the micro-mechanisms, since UN security forces are highly differentiated by troop contributing country, and independently conscripted and trained. Lastly, the mechanisms discussed here are broadly applicable, and may also be transferable to military effectiveness in inter-state conflict.

My findings have multiple implications. They shed light on the relationship between differentiation and COIN effectiveness, and provide a further step towards resolving the discrepancy and contradicting assumptions in the existing literature on the relationship between counterbalancing, differentiation, and COIN effectiveness (e.g. J. M. Powell 2012b versus Pilster, Böhmelt and Tago 2016). This constitutes an essential contribution to the core of civil-military studies. The findings also bear practical policy implications, highlighting the importance of considering a military's force and command structures beyond the context of coup-proofing, potentially aiding in the future building of COIN capacity. By identifying and developing the concept of differentiation, this article has problematised core literature in the field, and sets a precedence for future work on military structure.

2.6 Online appendix

2.6.1 Appendix A: Descriptive statistics

Figure 3: Index adjustment



Three exemplary cases, Russia in 2001, Mali in 2015, and Guatemala in 1986, are highlighted in red to better illustrate the adjustments across and within categories of j.

2.6.2 Appendix B: Supplementary estimation results

Table 3 contains alternative model specifications for exploring the results' model dependence and comparative fit. Model S1 considers the additional control variables population size (logged) and ethnic fractionalisation, both retrieved from Carey, Colaresi and Mitchell (2016), which may both arguably influence conflict outcomes and the military's structure. Conflict duration is added separately in Model S2 due to the risk of post-treatment bias. As Callwell (1996, 129) puts it however: 'In campaigns where

Table 3: JAGS ordered logistic regression

	Dependent variable:			
	Conflict Outcome			
	(S1)	(S2)	(S3)	(*null*)
j_1: Rebel win | Draw	−3.664	−5.122	−3.561	−5.090
	(22.73%)	(14.92%)	(24.07%)	(6.82%)
j_2: Draw | Gov. win	−0.055	−1.467	0.153	−1.799
	(50.25%)	(38.3%)	(50.58%)	(29.95%)
Differentiation Index	−0.931	−0.961	−1.037	
	(1.57%)	(1.72%)	(1.35%)	
Coup Risk	1.368	1.518	1.485	−0.041
	(90.72%)	(92.58%)	(92.73%)	(40.67%)
Coup Risk*Diff.	−0.335	−0.380	−0.344	
	(9.98%)	(7.65%)	(8.62%)	
GDP/c	−0.366	−0.504	−0.385	−0.423
	(22.15%)	(16.53%)	(23.77%)	(16.5%)
Polity2	−0.014	−0.011	−0.02	−0.04
	(39.98%)	(42.72%)	(36.38%)	(20.9%)
Military Exp.	0.699	0.726	0.638	0.363
	(89%)	(88.87%)	(85.1%)	(75.82%)
Oil/c	1.313	1.425	1.505	1.135
	(97.65%)	(98.1%)	(98.75%)	(96.82%)
Population Size	0.386	0.362	0.395	
	(90.93%)	(87.58%)	(90.47%)	
Ethnic Frac.	1.293	1.232	1.403	
	(88.7%)	(86.72%)	(90.2%)	
Conflict Duration		−0.005	0.015	
		(47.87%)	(56.25%)	
Polity2^2			0.015	
			(90%)	
Observations	73	73	73	73
Mean deviance	133.24	134.78	134.18	134.35

Posterior means reported. Percent positive in parentheses

the hostile tactics have been limited to surprises on a small scale and to ambuscades, and which have drifted on in a desultory fashion for long periods, the want of prompt energetic action on the part of the regular troops has generally been mainly responsible for their failure.' In addition to its effect on conflict outcome, longer conflict duration may cause an adaptation of the security apparatus, which in combination renders it a relevant confounder. Finally, Model S3 offers a quadratic term of Polity2. All of these additional variables do not attenuate the effect of either differentiation or the interaction term, which may increase confidence in the results.

The decrease in the deviance from the *null* model to the main model is negligible, meaning that differentiation and its interaction with coup risk only marginally contribute to the model's fit. However, expecting major changes in model fit may be demanding considering the small sample and limited variation in the slow-moving variables – a fact that is signified by the small degree to which the mean deviance is affected by any other changes to the model (S1-S3). For comparison, the mean deviance of a constant-only model is 134.71. These negligible differences in model fit are equally reflected in the models' AIC and BIC.

3 Jumping on the bandwagon:

Differentiation and security defection during conflict

When confronted with mass uprisings, governments deploy their security forces for crowd control or repression. However, sometimes security agencies choose to side with the opposition movement. Recent work shows that "fragmentation" contributes to defection: fragmenting the security forces into parallel units leads to oversight problems and grievances among soldiers, which raises the risk of members of the security forces defecting to the opposition movement. However, I find that the effect on defection is strongly moderated by the circumstances under which states choose to fragment their military: fragmentation for the purpose of security specialisation, called "differentiation", even decreases its risk. Employing Bayesian multilevel modelling, the findings corroborate this distinction. The study contributes to the fundamental discussion on civil-military relations, shedding light on why some conflict situations see security defections while others do not. Understanding this phenomenon is a pivotal element to explaining how conflicts develop, escalate, and end.

Keywords: defection, fragmentation, military effectiveness, civil-military relations

Article and data published at:

https://journals.sagepub.com/doi/10.1177/0022002720904763

3.1 Introduction

"Loyalty" is paramount for military success,[24] which is why insubordination and desertion are persecuted and severely punished – this was the case centuries ago, and remains true today (Dwyer 2015, 6). Despite harsh counter-measures, disobedience is a recurring phenomenon in all militaries, which is why an identification of the conditions under which soldiers are more likely to defect is as relevant as ever. In addition, these cases of secession cause more than "just" military ineffectiveness: they pose a threat to the underlying command structure as a whole, and therefore to the fundamental civil-military relations within a state.

Especially governments that take extensive measures to ensure their military's allegiance ("coup-proofing") have to expect an increased likelihood of military defections during (non-)violent uprisings (Dahl 2016a; Lutscher 2016). One of these coup-proofing measures refers to a fragmentation of the security apparatus, and is called "counterbalancing": a government that fears a coup d'état to happen fragments its security forces into multiple independent command structures, hence if one unit turned rogue, other loyal units could be deployed to "counterbalance" it. In their pioneering work, Dahl (2016a) and Lutscher (2016) find that such counterbalancing induces frustration, oversight problems, and ineffectiveness in the security apparatus, leading to more defections. I extend their work by arguing that features like frustration and ineffectiveness stem from the government's desire to coup-proof, and not from "fragmentation" itself. There are many countries maintaining a highly fragmented security apparatus *without* any coup-proofing intentions. Such fragmentation instead aims at a *specialisation* of individual force segments in different security tasks (e.g. counterinsurgency versus conventional warfare), and is herein called "differentiation". In contrast to counterbalancing, differentiation bears positive implications for effectiveness and capacity (Dworschak 2019). Consequently, while previous studies shed light on the link between coup-proofing measures and defection, I inquire 'how does military differentiation influence the likelihood of defections during (non-)violent uprisings?'.

This paper proceeds as follows. First, I conceptualise military defection and review different causes. Second, I explain the two types of fragmentation, "counterbalancing" and "differentiation", and develop an argument explaining how the latter yields opposite implications for defection during internal conflict. Security forces that are divided into multiple parallel units, each specifically trained and

24. See Chenoweth and Stephan (2013), Nepstad (2011), Stephan and Chenoweth (2008), Peaty (2002) and Brinton (1938).

equipped to conduct a distinct security task, are expected to experience a lower likelihood of defection. This is because grievances among soldiers are lowered, and civil unrest is effectively countered without it posing a threat to the government's survival. Third, the hypothesis that higher differentiation leads to a lower propensity of defection is evaluated using a Bayesian multilevel logistic analysis, which yields results in line with my theory. Subsequent robustness checks provide further corroboration. I conclude with a summary of my findings and avenues for further research.

3.2 Defection in the security apparatus

Military disobedience has many faces, and may refer to defections, coups, mutinies, or desertions. Past research on either of them has often failed to acknowledge previous work on any of the other concepts, despite them being so closely related. In this subsection I provide a short overview of the conceptual differences in order to unambiguously demarcate the definition of defection as it is used for this paper, and to start bridging the gap between the different strands.

Studies on desertion mostly offer in-depth descriptions of historical cases. In his monograph on desertion during World War I, Chen (2016, 1) details that "[u]nlike cowardice (if a soldier turns his back on the enemy and runs away), desertion was *subsequent* absence from the scene of danger with intent of avoiding danger" (emphasis original; definition based on Peaty (1999, 199)). Mutinies on the other hand occur both during conflict and peacetime, and have seen a recent surge in contemporary case studies and quantitative investigations. The concept represents "collective revolts by soldiers that are not focused on taking control of the state[, ... but instead] to convey soldiers' grievances." (Dwyer 2017, 4). Emphasising some of its possible dimensions, Rose (1982, 561) defines a mutiny as "an act of collective insubordination, [...] both active and passive[, ...] with or without violence." The key difference to deserters is that mutineers normally *stay* within the state's security apparatus and do not seek to evade their superiors, but instead aim at openly confronting them with certain demands, oftentimes trying to draw as much (public) attention to their cause as possible (Dwyer 2015, 7). This may lead, for example, to soldiers seizing their barracks or occupying a city (see e.g. Burke 2017) in order to voice their grievances about payment issues (Schiel, Faulkner and Powell 2017). Finally, individuals can desert unilaterally, while mutinies are collective endeavours that require coordination between soldiers.

While defection is sometimes used synonymously with desertion (see e.g. McLauchlin 2015), I employ a more narrow approach and build on the operationalisation by Dahl (2016a) and Lutscher (2016). Specifically, I understand defection to be a mixture of both desertion and mutiny: defecting soldiers *leave* the state military, in order to go into *open rebellion* against it. During civil conflict, this means that part of the security apparatus breaks away from the government and actively sides with the opposition movement ("side-switching"; c.f. Dahl 2016a; Dahl 2016b; Lutscher 2016).[25] Prime examples constitute militaries siding with the protesters in Egypt and Tunisia during the Arab Spring in 2011 (Brooks 2017). In order to address the imminent research gap and properly connect to the threads of Dahl (2016a) and Lutscher (2016), this paper solely applies this narrow definition of "defection".[26]

Coups d'état constitute the largest rupture in civil-military relations, and may be seen as "severe forms" of defections.[27] However upon closer inspection, their similarity proves superficial. First, defections are different from coups in that they require much less coordination (Dahl 2016a, 45-47). While coups have to be orchestrated and are directed at the centre of power, defections often happen in the periphery and can be perpetrated by individuals. Second, defections require an active opposition to which the soldier(s) can defect to, while coups may happen outside of active conflict environments (Lutscher 2016).[28] Third, building on the distinction drawn above, defecting soldiers leave the military, while coup participants often wish to conserve the military's structure and keep (or increase) their ranks, "only" aiming for a seamless turnover in state leadership (Derpanopoulos et al. 2016; Thompson 1975, 485-486).

Following T. Lee (2009, 646-647), and reiterated by Lutscher (2016, 356), one important factor determining defection is bandwagoning: when segments of the military recognise that they are fighting on the "losing side", i.e. that their government will inevitably succumb to the domestic opposition (possibly paired with international pressure), the risk of defection surges. Similarly, Nepstad (2013) suggests that perceived fragility of the regime is a key driver for disobedience and mutiny: an unwillingness of international allies to help a struggling government may signal security forces that their principal

25. The definition brought forward by Gates (2002) is the other way around, labelling NAVCO's defection as "desertion", while his term "defection" instead refers to just not performing an assigned task; a jumble that may be due to the game theoretic cooperation-defection dichotomy.

26. "Staying quartered" (Pion-Berlin, Esparza and Grisham 2014), also referred to as idleness, may be the result of either defection or mutiny. Soldiers stay in their barracks or refuse to engage protesters as a signal of open dissent. However, this may either imply a divorce from state leadership (defection), or a means of communicating dissatisfaction with their military leadership without seriously intending to risk the government's survival (mutiny). In times of conflict however, idleness may be equated with defection in most cases.

27. Many thanks to the anonymous reviewer who pointed this out.

28. With some important exceptions; c.f. (Bell and Sudduth 2017).

will not prevail. An example of this is Gorbachev's decision to stay idle over the 1989 uprisings in East Germany, spurring large-scale defections (Nepstad 2013, 340). However, bandwagoning is not only a substantial explanans for individual-level defections due to security forces avoiding to fight a losing battle, but also because of their organisational preference for cohesion: Pion-Berlin, Esparza and Grisham (2014) indicate that military leadership eschews ordering deployments that may be divisive among soldiers, and that could lead individual units to break away from the force. Uncertainty over the regime's survival, as well as suppressing (especially non-violent) insurrections, critically endanger such internal cohesion, and increase group-level defections (Lutscher 2016, 356; Pion-Berlin, Esparza and Grisham 2014, 233-235).

Finally, Dahl (2016a, 2016b) shows how major incidences of military personnel siding with the opposition movement are rendered more likely when a state employs counterbalancing. "Counterbalancing" is a tool that is employed in countries which are at risk of experiencing a coup d'état: it refers to the purposeful fragmentation of the military so that power is not concentrated within a single command structure, and individual segments can be used against each other in case of one staging a coup (Quinlivan 1999, 141; Belkin and Schofer 2003, 613; Belkin and Schofer 2005, 147-150). According to Dahl (2016a), counterbalancing fosters grievances in the security apparatus and impedes effective monitoring, both of which enable defections. Lutscher (2016) adds to these findings, showing that the relationship between counterbalancing and defection is U-shaped: while he agrees that low and high degrees of counterbalancing lead to an increased likelihood of defection, a medium level of counterbalancing however sees a decrease. This is because at a medium level of counterbalancing, its originally intended effects (preventing military segments to turn against the government) are most prominent, thus hampering defection (355-358).

3.3 Expanding the discussion: Military fragmentation

I propose a new approach to explain the occurrence of defections. Beginning at the literature's point of departure, I first review the implications of "counterbalancing" (cf. Dahl 2016a; Lutscher 2016), and then proceed to compare them to "pure" military fragmentation under no coup risk, which I call "differentiation".

3.3.1 Two types of fragmentation

There are two distinct types of military fragmentation: counterbalancing and differentiation. This difference becomes apparent when considering the fact that force fragmentation occurs in various states with both high and low coup risk, indicating that such fragmentation may not solely serve coup-proofing purposes. Instead of coup-proofing, a common reason for establishing multiple parallel militaries is specialised task-sharing: while one army may be trained and equipped for inter-state warfare, the other is employed for internal pacification. Although this difference in intentions for fractionalising the security apparatus may seem subtle, it has profound implications for military effectiveness (Dworschak 2019), and, as I argue in this paper, the likelihood of defection.

Previous research has considered the effect of counterbalancing on the likelihood of defection, while disregarding this other type of force fragmentation (Dahl 2016a, 2016b; Lutscher 2016). I seek to remedy this gap by hypothesising counterbalancing and differentiation to have contradicting influences on the likelihood of defection, leading to a unified framework of the relationship between fragmentation and defection during conflict.

Counterbalancing

Broadly speaking, a government faces two "extraordinary" (non-institutionalised) threats to its survival: on the one hand interventions by the elite, which are commonly referred to as coups d'état, and on the other hand popular uprisings, ranging from large-scale protests up to violent insurgencies (Roessler 2016; Svolik 2012, 3-10; Acemoglu and Robinson 2001). States are aware of these threats, and adapt their security apparatus accordingly.

Starting with the first type, especially authoritarian regimes face a higher risk of experiencing coup attempts (Böhmelt, Ruggeri and Pilster 2017, 225). In order to minimise this risk, they can implement various coup-proofing strategies, one of which is called counterbalancing: based on the idea "divide and conquer" (Belkin and Schofer 2005, 147-148), the state leadership divides its military into multiple sub-units with separate command structures, as to avoid concentrating too much power within few commanders. In addition, it raises associated coordination problems in case a military commander wants to topple the government, and it allows the leadership to deploy the remaining loyal parallel militar-

ies against the renegade unit (to "counter-balance" it).[29] Counterbalancing is generally accompanied by a set of other coup-proofing measures, which are outlined by Quinlivan (1999) in light of cases in Saudi Arabia, Syria, and Iraq: when in fear of coups, the leadership generally seeks to hamper forces' capacity and resolve, and tends to foster military promotions not based on merit, but instead on ethnic, religious, or family ties, in an attempt to secure their commanders' allegiance. Joint training exercises, another important feature for military effectiveness, are also discouraged, as they could serve to overcome the purposefully induced coordination problems.

In summary, governments that face a heightened coup risk perform coup-proofing, which includes counterbalancing.[30] However, as an externality of this behaviour, these governments may simultaneously render themselves more susceptible to the second type of threat: popular uprisings (Roessler 2016). As Biddle and Long (2004, 532) point out,[31] hampered resolve and training, as well as partisanship in the promotion of staff, lead to a decrease in human capital, military ineffectiveness, and generate frustration among many personnel – which is one of the reasons why coup-proofing is thought to increase the risk of defection (Dahl 2016a, 214). Another coup-proofing feature is the constant rotation of top officers, which aims at preventing "alternative power bases from developing", but again increases risks of partial defections due to impeded oversight (216).

Differentiation

In this article I seek to accommodate the fact that coup-proofing is not the sole reason for states to fragment their security apparatus, and that this difference is meaningful when estimating the likelihood of defections. The main alternative reason for governments to pursue military fragmentation, other than counterbalancing, is to achieve specialisation. For example, India has a highly fragmented security apparatus, but a low structural coup risk.[32] In this context, high fragmentation does not take place in order to forestall a coup, but to have multiple security forces trained and equipped in different tasks:

29. For an overview, and for further discussions on counterbalancing, see e.g. Belkin and Schofer (2003, 613), Belkin and Schofer (2005, 149-150), Pilster and Böhmelt (2011, 335-336), J. M. Powell (2012a, 40-41), and J. M. Powell (2012b, 1022-1023).

30. This assumption is not uncontested. Appreciating the military and government to be strategic actors, and subsuming the prohibitive effects of coup-proofing under the concept of coup risk, Sudduth (2017) shows that higher coup risk leads to a *decrease* in counterbalancing. "Coup risk" then refers to the predicted probability of observing a coup attempt. For the purpose of this study, I follow Belkin and Schofer (2003) in conceptualising coup risk as a latent variable capturing the underlying risk environment, and exogenising coup-proofing. This approach leads some countries to have a high coup risk despite not exhibiting any realised coups due to effective coup-proofing measures.

31. See also J. M. Powell (2012a, 140-145).

32. The coup risk of India has been continuously decreasing over the years. While its overall mean is within the first (lowest) quartile of the coup risk index, its number of parallel security forces at any point in time is the highest in the whole dataset. See the section "Research Design" for the data sources.

for one, India has a need to maintain conventional military capacity (large, mechanised forces) in light of the contestation with Pakistan, and to deter other potential inter-state challengers (Ladwig 2015; Kapur 2008). Additionally however, the Indian government faces several insurgencies, which generally require non-conventional, mobile units that are specially trained and equipped for asymmetric warfare (Pilster, Böhmelt and Tago 2016; Lyall and Wilson 2009; Dworschak 2019). Therefore, the fact that India maintains multiple parallel security forces is not a result of coup-proofing, but of task specialisation.

While there is a consensus that counterbalancing, and many of the above-mentioned coup-proofing measures, are detrimental to military effectiveness, motivation, and cohesion (Roessler 2016; J. M. Powell 2012a; Biddle and Long 2004), these negative implications may not be shared by the type of fragmentation as in the case of India. On the contrary, the argument I bring forward assumes that an employment of multiple parallel forces for the purpose of tailoring their training and equipment to fulfil specific security tasks necessarily leads to an increase in their effectiveness and endowment (cf. Pilster, Böhmelt and Tago 2016; Dworschak 2019). Such "pure" fragmentation that is distinct from counterbalancing, since it does not come with any other coup-proofing measures and does not aim at purposefully rendering the military *less* capable, is herein referred to as "differentiation".[33]

3.3.2 Differentiation and bandwagoning during conflict

In cases of defections during civil conflict, both types of "extraordinary" threats to regimes' survival fall together, one from the masses and one from the elites: a popular (non-)violent movement poses a bottom-up threat to the ruler, combined with parts of the government's security forces joining this uprising. Therefore, these cases constitute an especially interesting setting to compare the influences of differentiation and counterbalancing on defection, as each speaks to a different type of threat.

Security forces that are subjected to counterbalancing and other coup-proofing measures are less apt to confront an uprising, since governments facing high coup risk seek to hamper military capacity, allow less joint training, decrease human capital on both the officer and soldier levels, and impede communication and coordination (see e.g. J. M. Powell 2012a, 140-145; Biddle and Long 2004, 532; Huntington 1957, 82-85). Soldiers however generally prefer not to fight on the losing side (Lutscher 2016, 356; Nepstad 2013, 340; T. Lee 2009, 646-647). When soldiers begin to appreciate that their efforts are futile and that they will not be able to end a seemingly successful revolt, they will rather "jump on

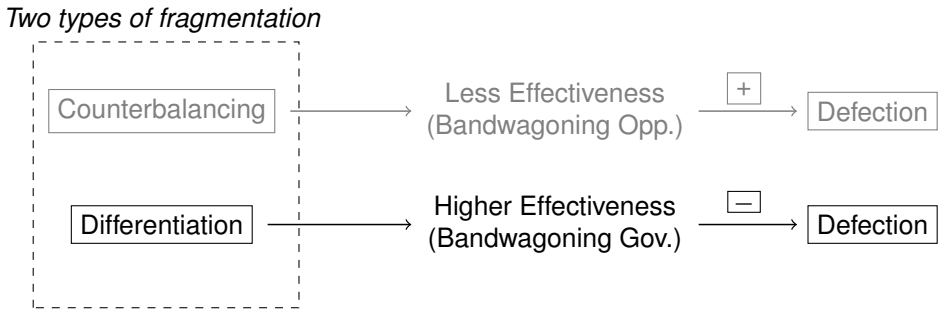
33. A term that was originally coined by Pilster, Böhmelt and Tago (2016), and conceptually developed by Dworschak (2019).

the bandwagon" and side with the opposition, instead of remaining with a government that they expect to eventually be overthrown. Drawing on the fact that differentiation is associated with more specialised training and equipment, more endowments, and *increased* military effectiveness, this would indicate that the bandwagoning effect becomes reversed: a differentiated military is better able to counter revolts and insurgencies. This stifles the opposition's chances to succeed, which in turn renders the security forces more prone to stay aligned with the government. Past literature mainly considers grievances and motivation as drivers of defection, but tends to neglect the relevance of human resources in the military, and its *ability* to confront an opposition movement. I argue that military ineffectiveness, or more specifically, being unfit to meet the challenge of an uprising, significantly adds to our understanding of the link between fragmentation and defection.

However, especially when facing nonviolent uprisings, the military's motivation to support the regime remains a crucial factor determining their loyalty (Brooks 2017). Counterbalancing leads to decreased endowments and dissatisfaction among military personnel, which lead to an increased risk of soldiers and units to defect to the opposition (Dahl 2016a, 2016b; Lutscher 2016). Based on the notion that differentiation may be associated with increased funds, better training and equipment, and an overall increased leverage, it follows that differentiation decreases the risk of soldiers defecting. In summary, although both counterbalancing and differentiation refer to a fragmentation of the state's security apparatus, they constitute very distinct concepts, and have contradicting implications for security forces' effectiveness and motivation. This results in opposite expectations regarding a force's propensity to defect during conflict. See Figure 4 on the following page for a visual summary.

One example that is particularly illustrative of this framework is the Tunisian army. President Habib Bourguiba, predecessor of Ben Ali, established civilian control as the central tenet of Tunisian civil-military relations: in this rapport, the military was fundamentally professional and "non-political" (Barany 2011, 31). During his time, the Tunisian government perceived itself to be at a low coup risk, and while its military was highly fragmented, this was not particularly directed at counterbalancing or other coup-proofing purposes. On the contrary, the army was very well funded, granted a high capacity, and its three well-trained parallel brigades were allowed to engage in joint exercises (Pachon 2014, 525). In addition, the Interior Ministry maintained a large, parallel security apparatus especially trained and equipped for internal pacification, and the army would only be called in when these militarised police

Figure 4: Theoretical overview



The greyed out part, linking counterbalancing to an increase in defections, signifies the macro relationship that has been tested in the past. The lower part corresponds to my contribution.

forces would become overwhelmed by an uprising (Barany 2011, 31). This professionalism and clear task-separation added to the general reservation that most militaries hold against "policing tasks",[34] and resulted in the Tunisian army resenting to be ordered to help quell the uprisings of 1978 and 1984 (31). Despite this "situational frustration" however, there were no accounts of any insubordination, as the army experienced no problems in countering the revolts and had an overall high motivation to keep Bourguiba in power due to their beneficial status (Pachon 2014, 521-523).[35]

In contrast to this, Ben Ali purposefully kept his military small and under-funded, and deprived it of essential resources (Brooks 2013, 208). Accordingly, the defections that have led to his eventual downfall in 2011 were committed by exactly those units that were least able to confront large-scale riots, and that were deployed *although* they lacked the skill, training, and resources to face widespread protests (Pachon 2014, 528-529). This was quickly followed by other forces joining the defectors or resorting to idleness, as there was little motivation to keep a ruler in power who systematically kept the military in a tight grip and marginalised (Brooks 2017, 7-8; Brooks 2013, 216). In summary, the security forces' ability and motivation to counter an uprising constituted a crucial determinant for their propensity to defect. Differentiating the security apparatus, and thus maintaining specialised and well-equipped

34. See Pion-Berlin, Esparza and Grisham (2014, 234-235).

35. Interestingly, and again speaking to the coup-civil war trap (cf. Roessler 2016), in the aftermath of the so-called "bread riots" of 1984, internal restructuring of the security forces to gain further efficiency has set the stage for Ben Ali's coup in 1987 (Gana 2013, 22). While this example may seem to imply a positive effect of differentiation on coup risk, my theoretical framework does not lend itself to make such inference: defections and coups differ in their underlying dynamics, and in how they are entwined with the two types of fragmentation. Any statement about the link between differentiation and coups would require a follow-up study.

forces that are apt and willing to engage the security threat at hand, seems to be paramount to military cohesion.

3.3.3 Hypothesis

I argue that the adverse effect of counterbalancing on defection is not inherent to fragmentation itself, but relates to the concomitants of coup-proofing. Shifting the theoretical focus on military effectiveness and bandwagoning, it becomes apparent that differentiation, i.e. fragmentation for the purpose of security specialisation instead of coup-proofing, may even lead to a decrease in defections due to the security apparatus being better able and more willing to overcome (non-)violent uprisings. This leads to the directional hypothesis below, which I quantitatively evaluate in the following section.[36] Note that, in line with my theory and conceptualisation, the analysis will interact a measure of fragmentation with structural coup risk, since the term "differentiation" implies fragmentation under low coup risk. This will be further explained in the operationalisation below.

H_1: Increased fragmentation of a state's security apparatus under low coup risk
leads to a lower likelihood of defection during conflict.

3.4 Research design and empirical evidence

3.4.1 Variables and data

Building on my earlier conceptualisation, I measure the dependent variable using a binary indicator of whether a security defection occurred in a given campaign-year (unit of analysis) from the NAVCO 2.0 data project (Chenoweth and Lewis 2013). "Security defection" refers to segments of the state security apparatus (elements of the military, paramilitaries, or police) switching sides and openly supporting the (violent or nonviolent) opposition campaign. Often, these cases involve military leadership, however it may also be lower ranks or the police siding with the opposition. Sometimes multiple defections may happen over the course of one campaign. See Table 4 on the next page for summary statistics.

36. A hypothesis on counterbalancing is not included, since these findings can be reviewed in the comprehensive studies by Dahl (2016a) and Lutscher (2016). A successful replication of their analyses was possible.

Table 4: Data summary

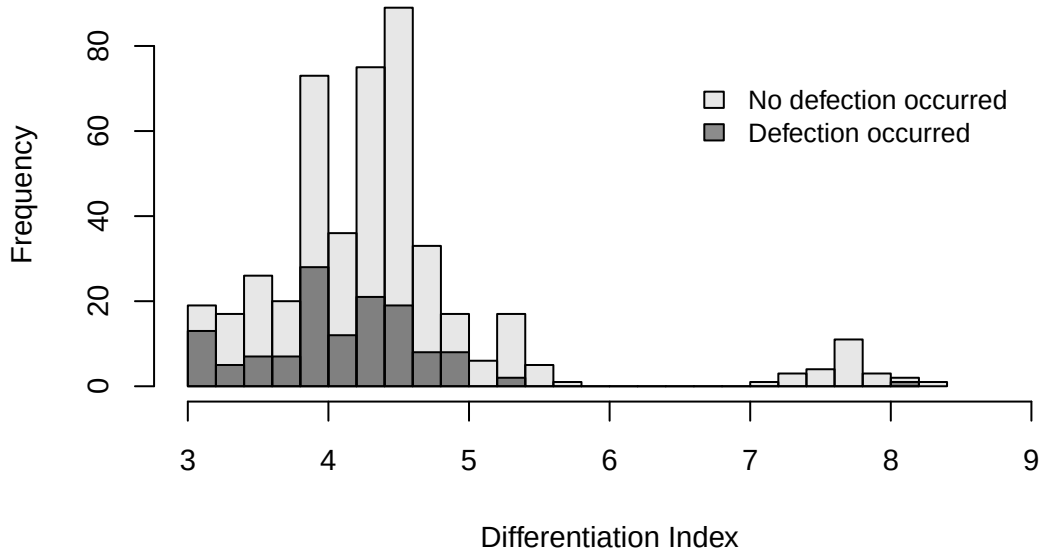
	Min.	Median	Mean	Max.	Data Type	NA's
Dependent Variable						
Security Defection	0	0	0.22	1	binary	336
Explanatory Variables						
Differentiation	3	4.31	4.26	8.21	ratio	34
Country-level Controls						
Coup Risk	-4.50	-0.32	0.05	7.68	ratio	151
Polity2	-9	0	0.30	10	ordinal	12
GDP/c	5.03	8.15	8.16	10.66	ratio, logged	69
Military Expenditure	0.00	1.34	1.38	3.43	ratio, logged	91
Campaign-level Controls						
Violent Campaign	0	0	0.22	1	binary	0
Regime Support	0	1	0.62	1	binary	146
Campaign Support	0	1	0.54	1	binary	56

928 total obs. (144 campaigns, 91 countries, years 1981-2006)

Based on Pilster, Böhmelt and Tago (2016) and Dworschak (2019), fragmentation of the security apparatus generally refers to the number and ratio of rivalling ground-combat compatible[37] security organisations in a state. These include for example the regular army, marines, border police, and other kinds of paramilitaries ("auxiliary forces"; cf. Böhmelt and Clayton 2018, 221). In line with my conceptualisation, and based on the distinction between counterbalancing and differentiation explained earlier, I adopt the adjusted differentiation index proposed by Dworschak (2019): $2\sqrt{j} + \Sigma_j s_{jit}^2$. Here, j refers to the number of parallel military units per country-year observation (i and t), and s constitutes each unit's share of overall military personnel per country-year. A higher index value thus indicates higher differentiation, with 3 being the baseline (no differentiation). For information on the numbers and sizes of security organisations, I draw on the raw data of Pilster and Böhmelt (2011) and Pilster, Böhmelt and Tago (2016), which are based on the Military Balance dataset of the International Institute for Strategic Studies (2016). See Figure 5 on the following page for a histogram on the distribution of security defections across different values of differentiation. The leverage points (differentiation index is seven or higher) pertain exclusively to India. To ensure that India is not the sole driver of the effect, an analysis that omits India is included in the appendix (see Table 8 on page 69).

37. In other words, all forces that are not purely naval or airborne.

Figure 5: Security defection across differentiation



Considering the theoretical expectation that differentiation and counterbalancing yield opposite effects on security defections, the crucial conceptual distinction between these two types of fragmentation also requires to be reflected empirically. This is accommodated in two steps. First, counterbalancing indices normally reward equal force sizes, meaning that countries with one small unit and one large unit receiver a lower value on the counterbalancing scale than countries with two equally-sized security organisations (Pilster, Böhmelt and Tago 2016; Pilster and Böhmelt 2011). This logic is built on the notion by Quinlivan (1999, 141-142), emphasising that for effective counterbalancing, the security forces must be “parallel militaries” instead of “paramilitaries”, referring to the necessity of them being “large enough” and equally well equipped in order to pose a credible threat to a deserting unit.[38] Using this theoretical approach, I turn the index around to instead reward unequal sizes: the existence of one small force next to a significantly larger one is probably not indicative of counterbalancing, as the small unit could not provide enough firepower for posing a credible counterweight. However, it may very well imply a

38. This approach came under increased scrutiny, and newer versions of counterbalancing indices do not take relative force sizes into account anymore (De Bruin 2018). Depending on the degree to which the comparison of relative force sizes is unreliable, I expect measurement error to attenuate the effects and render my estimates more conservative. In addition, this renders the second step (interaction) more relevant.

small strike force specialised in a security task other than conventional warfare. Comparing the conventional counterbalancing index with my adjusted differentiation index, they yield a correlation of $r = 0.54$. This sizeable correlation is not surprising considering that both counterbalancing and differentiation are measures of fragmentation, and I must expect the differentiation index to still be sensitive to different levels of counterbalancing.

Therefore, second, I interact the differentiation index with a measure of structural coup risk (Dworschak 2019; J. Powell 2019). The coup risk index by Belkin and Schofer (2003) captures the underlying propensity of a coup d'état happening (see also Belkin and Schofer 2005; Pilster and Böhmelt 2011; Carey, Colaresi and Mitchell 2016). It consists of measures of the strength of civil society, regime legitimacy, and recent coups (Belkin and Schofer 2003). The interaction allows the effect of the differentiation index on the propensity of security defection to vary across different levels of coup risk. This is useful because a high differentiation index value in a country with very *low* coup risk may safely be assumed to capture "pure" differentiation – which means that in accordance with my hypothesis above, under low coup risk I would expect a negative index slope. In contrast, a high index value in a country with a *high* coup risk might partly be driven by counterbalancing, and thus I expect the slope of the index in such strata to be attenuated or even turn positive (Dworschak 2019).

India constitutes an exemplary case that illustrates the substantial usefulness of drawing an empirical distinction between counterbalancing and differentiation. Despite multiple campaign involvements, India has not seen any security defections during the 1990s. However, past models by Dahl (2016b), assuming fragmentation to equate counterbalancing, would wrongly predict a high risk of defection. India's security apparatus was (and is) highly fragmented, leading to a high counterbalancing index: its mean counterbalancing value during the 90s is within the highest quintile of the overall distribution. I argue that India's fragmentation however is not aimed at hampering military resolve – on the contrary, it is the result of task separation among their armed forces, enabling specialised training and adaptation to local particularities (Dworschak 2019; Pilster, Böhmelt and Tago 2016). This is empirically signified by *a combination* of a high differentiation index (India is the most extreme case in the dataset) and a low coup risk. In conclusion, while India during the 90s shows high values for both counterbalancing and differentiation, the interaction with coup risk enables me to tell the two concepts apart, and

leads to predictions that are closer to reality. The same increase in accuracy may be expected for many other cases, including authoritarian regimes (cf. Lutscher 2016).

I employ several other control variables to further reduce the risk of bias. On the country level, a state's regime type, overall wealth, and military expenditure might profoundly influence civil-military relations both in terms of the propensity of security defection and the structure of the military. The first two are measured using GDP per capita and the country's Polity2 score, which are drawn from the replication data of Carey, Colaresi and Mitchell (2016). Military expenditure is provided by SIPRI (2017).

Another source of endogeneity that challenges my results' validity is the general threat environment underlying a campaign. Differences in the population's antagonisms and ease of mobilisation, as well as the external support structure for both warring parties, might lead the state to adapt its security apparatus accordingly from the outset, and to be fundamentally more (or less) prone to experience security defections. I address this omitted variable bias by including NAVCO 2.0 covariates on the campaign level that may serve to jointly proxy this latent dimension. These include whether a campaign was violent or non-violent,[39] and whether the regime or the opposition movement relied on external support (Chenoweth and Lewis 2013).[40]

3.4.2 Method

I employ a Bayesian two-level logistic regression with random intercepts and imputed missings. I briefly address each of these choices subsequently. First, considering the binary nature of my dependent variable, convention suggests the use of a logit or probit link. I choose the former due to a better interpretability of the coefficient estimates, although this has no substantial influence on the results. Second, since my analysis covers multiple countries from different regions around the world, I assume unobserved unit heterogeneity. This could lead to different base propensities for security defection, and to non-spherical errors. A multi-level model with random intercepts on the country level (normally distributed) helps to increase efficiency. While a correlation between the unit-level heterogeneity and my covariates is probable, partialling this out using fixed effects is not feasible due to low within-variation. For the results presentation, I average over all intercepts.[41]

39. Separate sub-sample analyses of violent and nonviolent campaigns yielded similar results, increasing confidence in the theory that division of labour benefits forces across these different security tasks, and that the effects are not driven solely by one type of campaign.

40. Assessing model dependence, the results also remain robust to an omission of these controls (cf. online appendix, Table 6 on page 66).

41. Due to my distributional choice, there is no substantial difference between the mean and the median intercept.

Third, 29% of the original data is missing. This is mainly introduced by the merging process, since the datasets substantially differ in their overall temporal coverage: the dataset by Carey, Colaresi and Mitchell (2016) starts in 1981, while the fragmentation data goes back to 1970, and NAVCO 2.0 even to 1945. Any kind of imputation over such large time periods seems questionable, which leads me to restrict the final data frame to only merge-able years and countries (corresponds to Table 4 on page 52). This final data includes 9% unobserved values, some of which still originate from a lack of spatial and temporal congruency, while others were coded as missing in the original datasets. I choose to impute these 9%, since the missings are scattered across different variables (cf. Figure 8 on page 63 in the appendix), and "standard" list-wise deletion would lead to a severe information loss (decrease from 928 to 415 observations; Little and Rubin 2002, 53-54).

Since it is not obvious what may have driven certain values to be coded as "missing", and it seems to be non-systematic, I assume they originate from the same data-generating process as the observed values (i.e., I conceptualise them to be "missing completely at random", MCAR; ibid., 11-13).[42] First, I reduce overall missingness to 8% by employing a combination of "last observation carried forward" (LOCF) and "next observation carried backward" (NOCB) for all slow-moving variables (Differentiation, Coup Risk, Polity2, GDP/c, and Military Expenditure). Second, I allow JAGS to impute the remaining missing values of the covariates based on approximate observed empirical distributions of my variables.[43]

I prefer Bayesian inference over a frequentist approach for multiple reasons. Firstly, multi-level modelling is an intrinsic feature of Bayesian (McElreath 2015, 14). While random effects models would not converge using standard procedures due to a lack of within-variation, this does not pose a problem with JAGS using (non-informative) priors. Secondly, in the Bayesian framework the missings in my dependent variable are not merely imputed, but are estimated as parameters and add to the overall uncertainty (Little and Rubin 2002, 200-203). Finally, Bayesian inference allows for a sensible probabilistic interpretation of the results (McElreath 2015, 4-17; P. M. Lee 2012, 139-140; Gill 1999).

42. This is a very strong assumption. However, the single imputation methods that I describe here yield similar results to multiple imputation based on observables, so while assuming MCAR versus MAR eases the analysis, it does not seem to bias the results.

43. The method by which I impute has no substantial effect on the results. Using the aforementioned multiple imputation, I generated 200 alternative datasets, and their average effect sizes remain similar compared to those presented here. Not imputing at all naturally increases the uncertainty around the point estimates.

Table 5: JAGS logistic analysis

	Post. means	Post. SD	Percent positive
Differentiation	−0.581	0.28	1.88%
Coup Risk	−1.089	0.69	4.92%
Diff.∗Coup Risk	0.227	0.16	92.74%
GDP/c	−0.587	0.34	3.69%
Polity2	−0.067	0.03	1.34%
Military Exp.	−0.969	0.43	1.20%
Violent Campaign	1.190	0.45	99.69%
Regime Support	0.193	0.38	69.14%
Campaign Support	1.431	0.46	99.97%
Intercept	*5.735*	*1.59*	

N chains: 3, Burn-in: 5000, Thinning: 200, Eff. sample/chain: 3000
N obs: 928

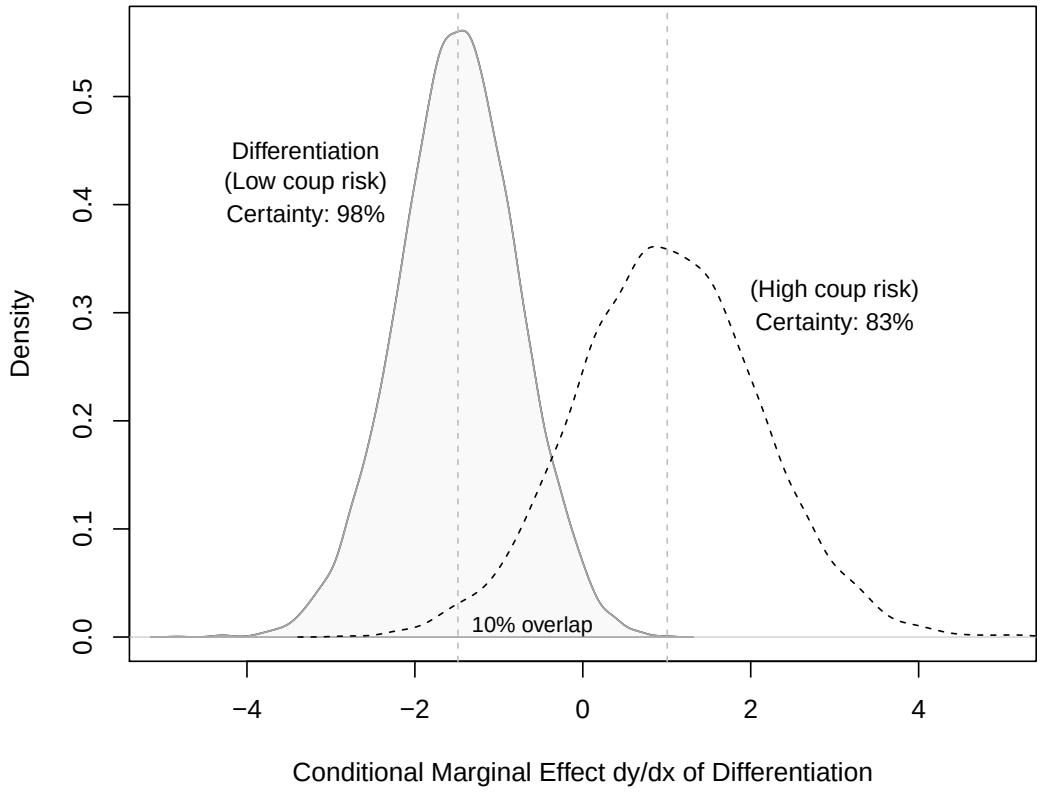
"Post. means" denotes the means of the posterior distributions. In a frequentist sense, these are the coefficients of the logistic regression. "Post. SD" are the standard deviations of the posterior distributions. "Percent positive" facilitates the interpretation of the estimates' (un)certainty by showing their probability of being larger than zero. The intercept reports the mean and standard deviation of the 91 countries' posterior means.

3.4.3 Results

The main findings are listed in Table 5. As expected, the effect of differentiation on security defection is on average negative. In turn, the coefficient of the interaction term is positive: this means that the effect of the differentiation index is strongly negative under low coup risk, but turns positive under high coup risk. This is in line with the aforementioned confounding influence of coup-proofing, and corroborates my theory.

In terms of uncertainty, these effects signify relatively low dispersion (high "statistical significance"). The probability that the differentiation index is indeed negative is about 98%, while the interaction term has a 93% probability of being positive. This means that based on these data, it is highly unlikely that these directional effects occurred by chance alone. Note however that even if the differentiation coefficient was centred on zero, this would not be indicative of the effects *within* the subsamples of high and low coup risk. Therefore, turning to the quantities of interest, Figure 6 on the following page shows the conditional marginal effects of differentiation. The density plot of the conditional posterior distribution

Figure 6: Marginal effects of differentiation on defection



Bandwidths = 0.1034 and 0.1596, n = 9000. Densities represent the marginal effect of fragmentation at coup risk values of -4 (Differentiation) and 7 ("Counterbalancing").

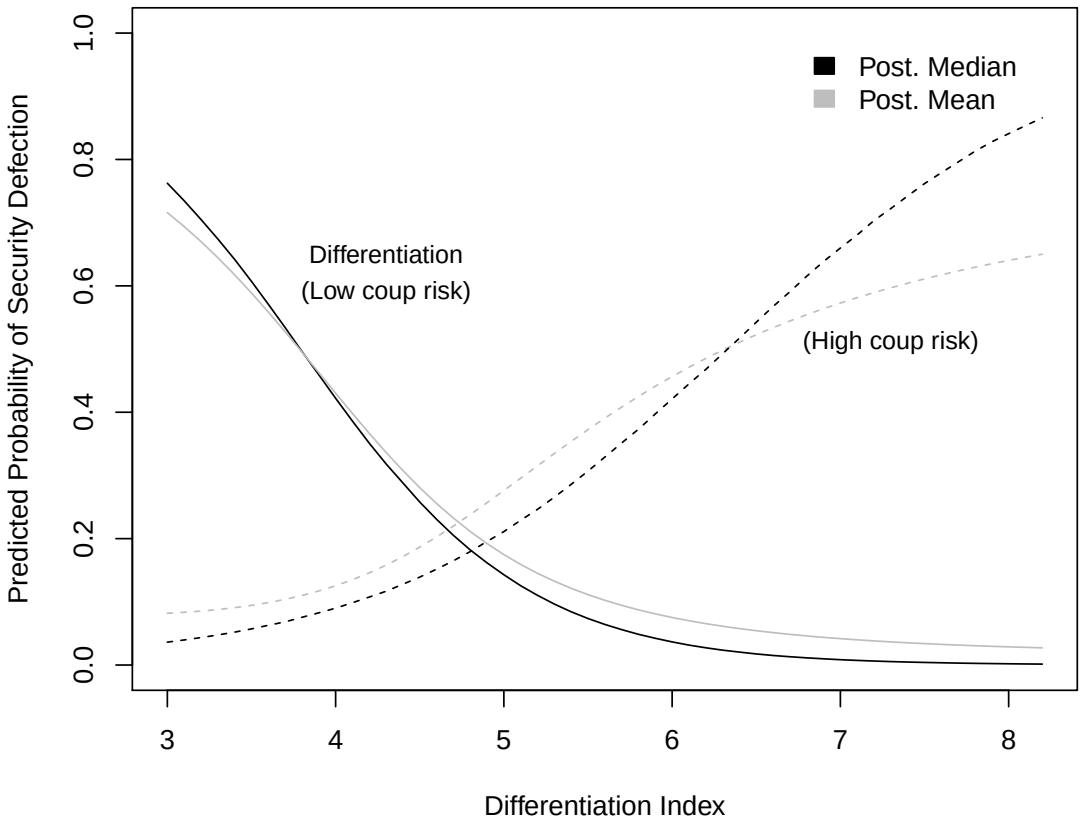
The effect of differentiation has a probability of 98% to be below zero. The effect of "counterbalancing" has a probability of 83% to be above zero. Their joint overlap is 10%.

of differentiation under low coup risk (minimum integer, -4) signifies a strong negative marginal effect. Under high coup risk (maximum integer, 7), the effect of differentiation moves much further to the right (more positive). I conclude that both a) the negative effect of differentiation on defection, as well as b) the difference of the impact of differentiation under low versus high coup risk, have a high probability of representing a systematic pattern.

For a more intuitive visualisation of the main effect, see Figure 7 on the next page.[44] The non-dotted lines correspond to the predicted probability of defection across different values of differentiation (low coup risk).[45] Assuming an average country (all other covariates are held at their median), it is apparent that differentiation strongly hampers the propensity for security forces to defect: while a state

44. Credible intervals are not plotted to avoid cluttering the graph. The results' joint uncertainty is fully visualised in Figure 6.
45. The two sets of closely aligned means and medians suggest symmetrical posterior distributions.

Figure 7: Differentiation and security defection



with minimal differentiation (3) has an absolute defection risk of 76%, an increase of differentiation to a medium level (5) already decreases this risk to 14%. Following my theory on differentiation as a means of force specialisation, a possible interpretation of these findings is that even a basic degree of task specialisation substantially increases effectiveness in internal pacification, and thus renders defections much less likely. In contrast, the set of dotted lines represent the same scenario, but under high coup risk. In this set-up, the degree of fragmentation that my differentiation index captures may rather be driven by coup-proofing intentions, and thus resembles counterbalancing. However, note that this is *not* the same as actual counterbalancing due to the adjustment in the index. Therefore, while the direction of the effect is highly suggestive and in accordance with the findings by Dahl (2016a) and Lutscher (2016), I refrain from interpreting the effect size as the actual impact of “counterbalancing”. These findings support my idea that the “intentions to fragment” the security apparatus (coup-proofing versus specialisation) matter substantially.

3.4.4 Robustness

The results are robust to a number of alternative model specifications, which are listed and explained in the online appendix. Table 6 on page 66 shows combinations of different model restrictions to assess model dependence and the predictive power of differentiation and the interaction term. The effect sizes remain similar, and differentiation and the interaction term increase predictive accuracy. Table 7 on page 67 continues with a model with all covariates lagged one year, and two models including additional controls (military size and autocratic regime types), with the results remaining robust. Table 8 on page 69 includes controls on more campaign features and population size, as well as a model with a quadratic term, a model that omits India, and a model in which missing values in the dependent variable are not imputed. The model that omits India experiences a slight attenuation in the effect size and certainty, which is further discussed in the appendix. Finally, Table 9 on page 71 tests additional empirical implications, offering alternative operationalisations of the intention to coup-proof. Specifically, I compare the effect of differentiation between democracies and autocracies, and between personalist and other kinds of autocracies, all of which are found to strongly differ in their coup risk. In combination with the index by Belkin and Schofer (2003), the results' consistency across these distinct measures for coup risk increases confidence in my findings. In terms of sampling, all parameters' MCMC chains in all models are well-mixed. See Figure 9 on page 64 in the appendix for example trace plots of the main parameters and GDP/c of the main model.

3.5 Conclusion

How does differentiation of the state's security apparatus influence the likelihood of defections during mass uprisings? In this study, I evaluate literature on militaries' loyalties and internal structure, and develop an argument surrounding military effectiveness and endowment. Past scholarship identifies fragmentation of the security apparatus as an important determinant for defections during mass uprisings. It is argued that parallel hierarchies in the military induce frustration and oversight problems, which in turn facilitate defections during conflict. I contend that these past findings apply to countries that engage in fragmentation for purposes of coup-proofing ("counterbalancing"), but not to countries that fragment their military for purposes of security specialisation ("differentiation"). Distinguishing between these two types of fragmentation is crucial, because it is associated with a stark difference in military effectiveness

and endowment: counterbalancing commonly correlates with low military capability, inferior equipment, marginalised power, promotions that are not based on merit, and low inter-unit communication and coordination. Differentiation on the other hand is aimed at specially trained and equipped forces, increased resources, and at effectively accommodating different security challenges.

Assuming that soldiers do not want to fight a losing battle, I claim that the military's ability to successfully confront a (non-)violent uprising is a strong negative predictor for defection. Therefore, countries that engage in differentiation are expected to experience less defection. I evaluate this hypothesis by quantitatively comparing the levels of differentiation and the likelihood of defections across countries. To have a reliable empirical distinction between counterbalancing and differentiation, I adjust the index according to my theory and perform the analysis conditional on countries' coup risk. Employing multilevel Bayesian logistic analysis, I find that countries that do not attempt to differentiate their security apparatus face an 80% probability of defection, while countries with maximum values of differentiation only experience a 5% probability of defection. The results are subject to fairly low uncertainty, and robust to different model specifications. I therefore conclude that differentiation indeed leads to a lower likelihood of security defection during (non-)violent conflict.

There are two main avenues for further research. First, merely controlling for confounding variables using macro data constitutes a rather weak identification strategy. Therefore, while my findings offer robust correlation with a high degree of generalisability, their internal validity requires future research on individual cases. Second, the extent to which data on horizontal fragmentation without explicit functional distinction can serve as a proxy for force specialisation is questionable, and I expect comparably low estimation efficiency and attenuation bias. Thus, the development and testing of more empirical implications of my theoretical model are worthwhile.

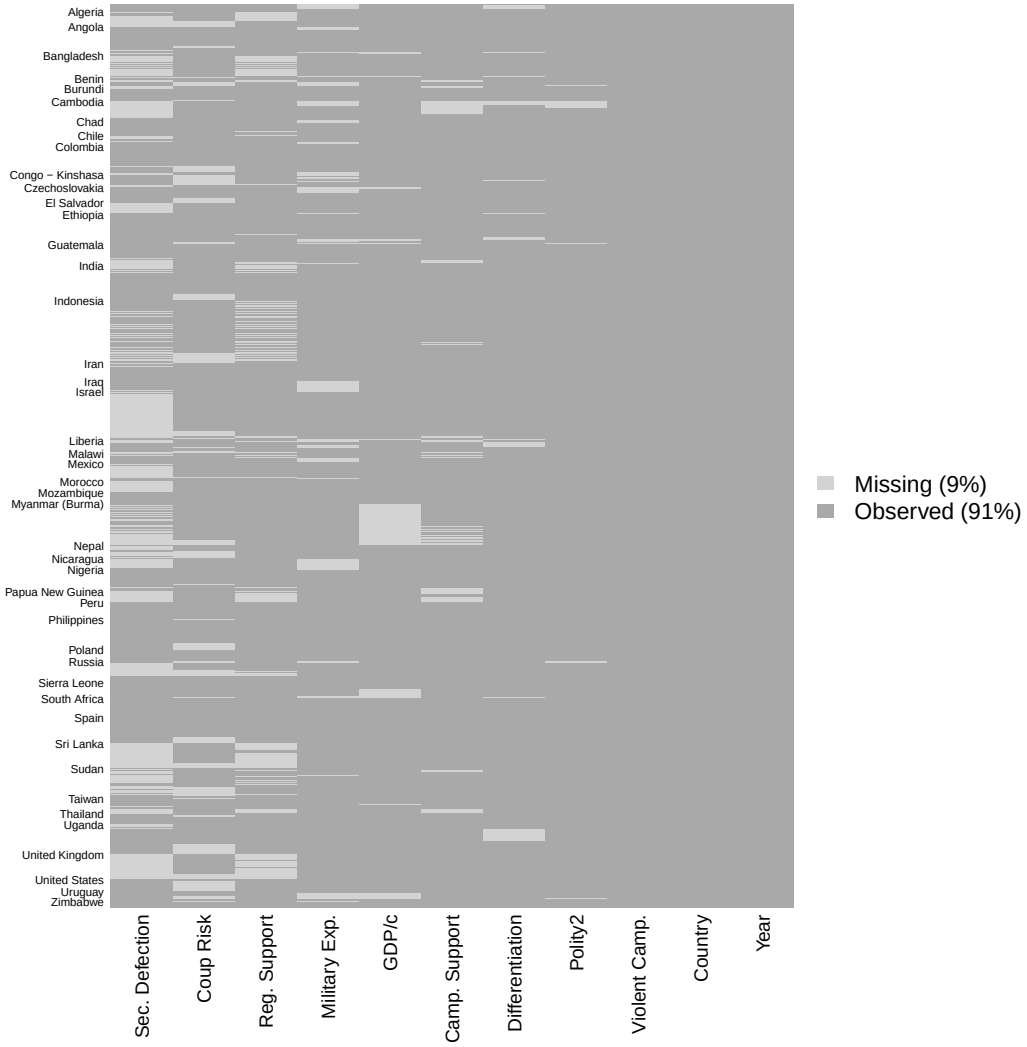
This study's contribution is three-fold: it is the first attempt to quantitatively substantiate the intuitive notion of bandwagoning in conflict, i.e. that soldiers do not want to fight on the losing side. The large effect size of differentiation underlines how "obvious" and unambiguous this mechanism is, rendering my findings a paramount corroboration of our intuitive understanding of military defections during conflict. Moreover, the article provides an important insight into the dichotomous nature of security force fragmentation, and into the opposing implications of differentiation versus counterbalancing. Lastly, it

yields policy-relevant information on civil-military relations and regime survival, highlighting the role of appropriate training for internal pacification and the double-edged effects of coup-proofing.

3.6 Online appendix

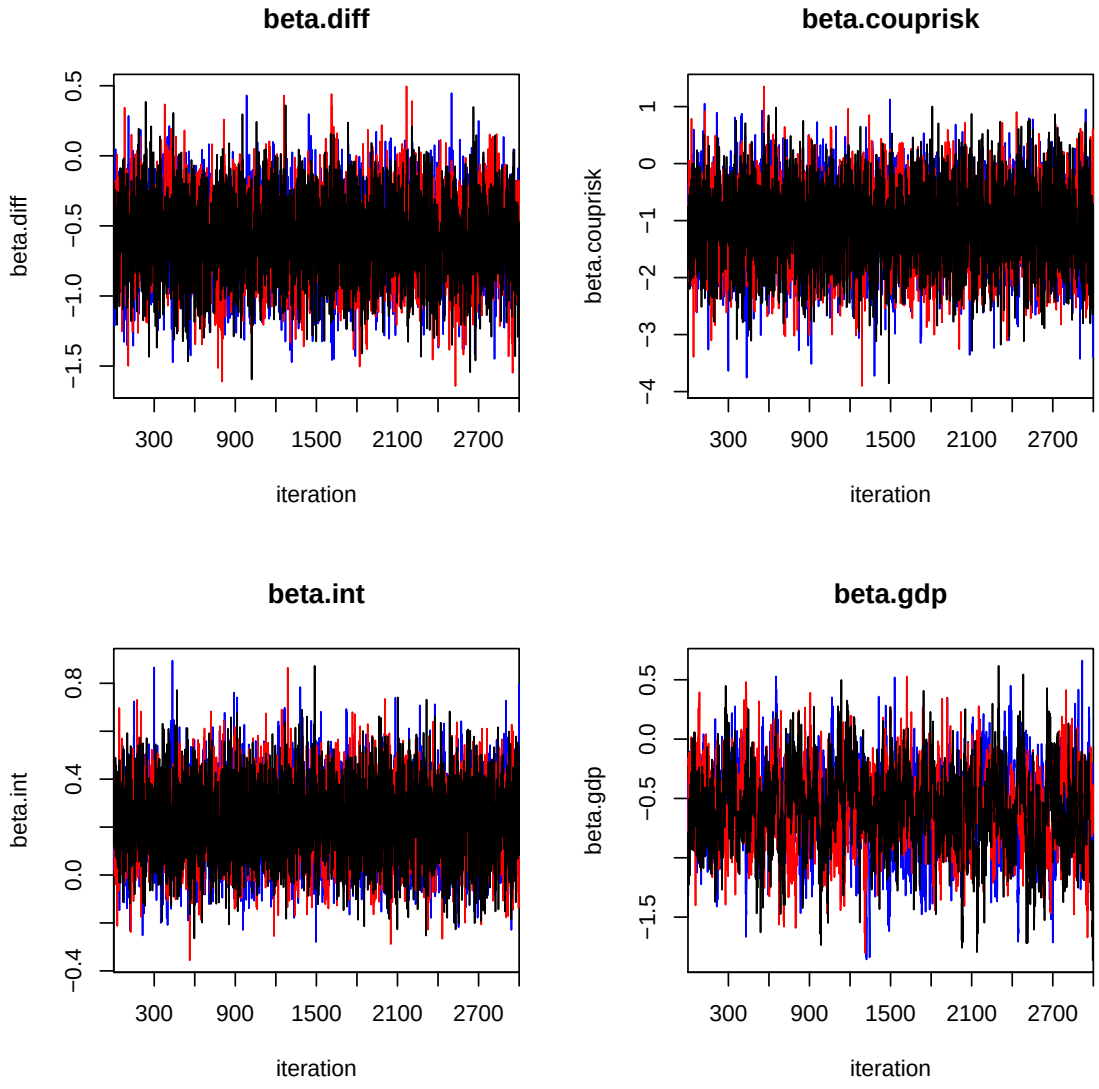
3.6.1 Appendix A: Descriptive statistics

Figure 8: Missingness plot



Dark areas show observed values, light areas show missing values. It is visible that most missings are induced by the dependent variable “Security Defection”. An analysis that does not impute these missing values is included in the appendix and shows similar results.

Figure 9: JAGS chains mixing



The chains of all parameter estimates are as well-mixed as the ones shown here for Differentiation (beta.diff), Coup Risk (beta.couprisk), and the interaction term (beta.int). The coefficient for GDP/c (beta.gdp) is included because it is the only sampling process that experienced major autocorrelation, which may come from GDP/c being slow-moving and yielding much less information than other variables. With 1,800,000 iterations (3 chains * 3000 stored samples * 200 thinning), I do not expect inconsistency to be a salient issue.

3.6.2 Appendix B: Supplementary results

Models I and II in Table 6 on the following page show different combinations of restrictions to illustrate the degree of model dependency. They suggest that the main effect's direction, magnitude, and uncertainty experience low model dependence. Restricting the coefficient of the interaction term (Models IV), and jointly restricting the coefficients of differentiation and the interaction term (Model III), leads to a small loss of fit compared to the main model. This is indicated in Figure 10, showing separation plots and Bier scores to assess predictive accuracy. In the separation plot of the main model, the realised defection events (black bars) are ordered more to the right, i.e. into the region of higher predicted probability, than the other two plots. Similarly, the Brier score (mean squared error) is slightly lower. However, these difference are only marginal.

Figure 10: Separation plots & Brier scores

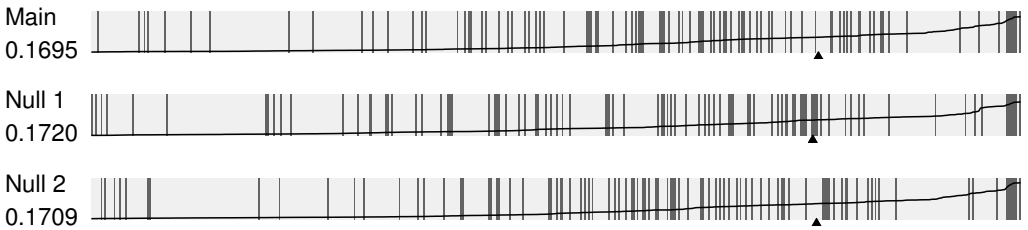


Table 7 on page 67 continues with three more models. In Model V, all covariates are lagged one year, in case the effect of differentiation is driven by defection in the same year leading to a restructuring of the security apparatus. The number of observations does not decrease due to the model's distribution imputations. Model VI includes a variable measuring the overall number of military ground personnel calculated from the IISS raw data described in the main body, since this naturally influences differentiation and possibly defection risk. Model VII additionally controls for different autocratic regime types, retrieved from Geddes, Wright and Frantz (2014), as they may show different defection propensities and inclinations to fragment their security apparatus. Model VIII includes cubic polynomials of the time since the last defection in order to account for time dependence (Carter and Signorino 2010). These alternative model specifications do not seem to impede the results' substantive significance or certainty.

Table 8 on page 69 shows four models. Model IX includes a set of additional controls for population size, and ethnic and religious diversity of the campaign. This is an attempt to partial out the

underlying threat environment in the state, as religious and ethnic cleavages (cf. Lipset and Rokkan 1967) may be associated with conflict propensity, severity, institutional cohesion, and the way in which forces are structured. However, note that these are merely binary indicators (NAVCO 2.0). Their inclusion does not infringe the results. Model X includes a quadratic term for differentiation as a comparison to the analysis on counterbalancing by Lutscher (2016). This induces severe multicollinearity, and while the size of the marginal effect remains about the same, uncertainty is inflated. Since there is no theoretical expectation for observing a curvilinear effect of differentiation, this is not troubling, and I refrain from further measures to address this issue (e.g. centering differentiation), leaving them to potential

Table 6: Alternative specifications I

	Dependent variable:			
	Security Defection			
	I: *no controls*	II: *country controls*	III: *null model 1*	IV: *null model 2*
Differentiation	−0.555	−0.540		−0.536
	(1.66%)	(2.63%)		(2.72%)
Coup Risk	−0.836	−1.067	−0.129	−0.144
	(8.6%)	(5.00%)	(5.21%)	(17.01%)
Diff∗Coup Risk	0.192	0.222		
	(91.17%)	(92.79%)		
GDP/c		−0.519	−0.652	−0.594
		(4.93%)	(2.61%)	(3.84%)
Polity2		−0.069	−0.066	−0.067
		(1.04%)	(1.92%)	(1.76%)
Military Exp.		−1.095	−0.939	−0.957
		(0.16%)	(1.21%)	(1.21%)
Violent Campaign			1.218	1.219
			(99.77%)	(99.73%)
Regime Support			0.099	0.147
			(60.48%)	(65.28%)
Campaign Support			1.406	1.404
			(99.97%)	(99.94%)
Intercept	*1.012*	*6.474*	*3.794*	*5.561*
	(1.33)	*(1.41)*	*(1.63)*	*(1.60)*
Observations	928	928	928	928
Note:	*Posterior means shown. Percent positive in brackets. Simulation & intercept specifications same as main model.*			

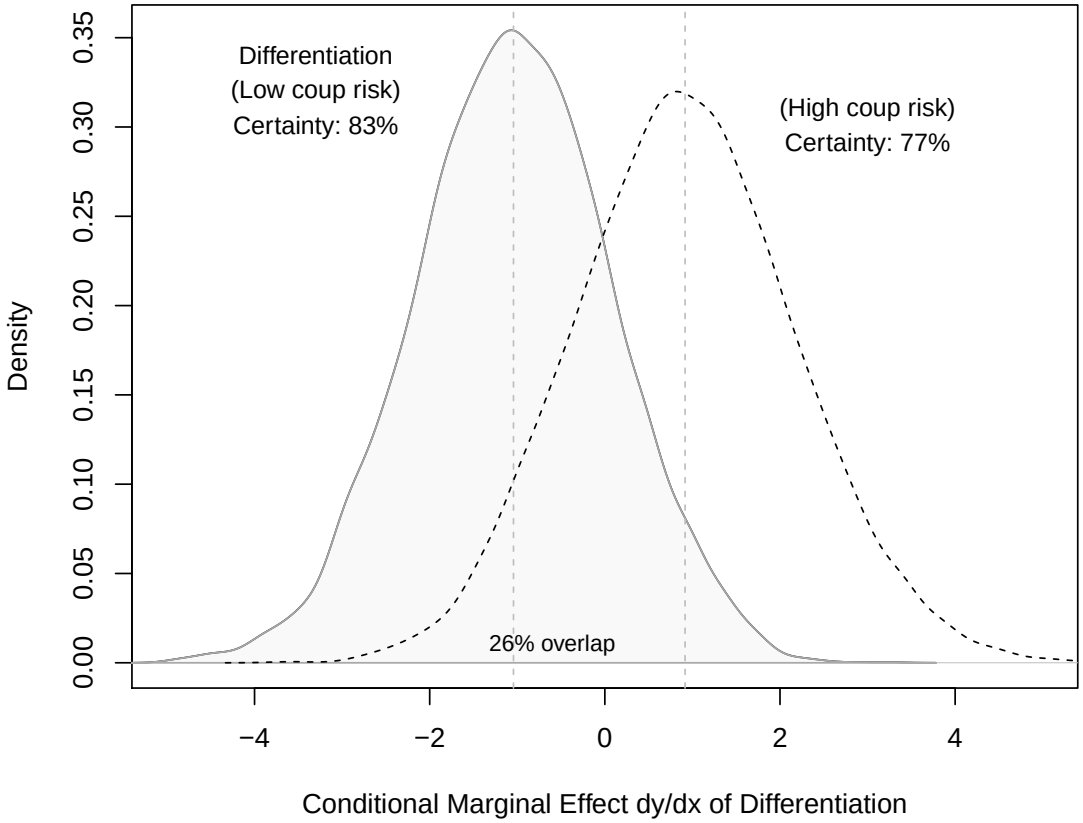
follow-up studies. Since India is the country showing the highest differentiation values in the sample, thus constituting a potential leverage case, Model XI omits the country. As anticipated given India's good

Table 7: Alternative specifications II

	Dependent variable:			
	Security Defection			
	V: *lagged covariates*	VI: *military size*	VII: *regime controls*	VIII: *time dependence*
Differentiation	−0.512 (10.14%)	−0.435 (8.40%)	−0.478 (5.9%)	−0.567 (2.2%)
Coup Risk	−1.621 (5.22%)	−1.250 (3.33%)	−1.097 (5.21%)	−1.124 (4.7%)
Diff∗Coup Risk	0.433 (97.33%)	0.261 (95.07%)	0.219 (91.78%)	0.228 (92.6%)
GDP/c	−0.876 (10.32%)	−0.520 (7.82%)	−0.459 (9.13%)	−0.606 (2.9%)
Polity2	−0.112 (0.98%)	−0.072 (1.03%)	−0.028 (23.42%)	−0.065 (2.0%)
Military Exp.	−3.001 (0.00%)	−1.027 (0.66%)	−0.879 (2.73%)	−0.969 (1.1%)
Violent Campaign	1.332 (96.38%)	1.203 (99.68%)	1.141 (99.54%)	1.113 (99.3%)
Regime Support	−0.364 (27.74%)	0.155 (65.24%)	0.246 (73.24%)	0.213 (71.4%)
Campaign Support	3.22 (100%)	1.472 (99.97%)	1.504 (99.99%)	1.426 (99.9%)
Military Size		−0.236 (10.70%)	−0.162 (19.37%)	
Military Aut.			1.095 (90.98%)	
Personal Aut.			1.119 (96.53%)	
Intercept	*8.854 (3.15)*	*7.393 (1.61)*	*5.546 (1.52)*	*6.187 (1.52)*
Time cubic polynomials	X	X	X	✓
Observations	928	928	928	928

Note: *Posterior means shown. Percent positive in brackets. Simulation & intercept specifications same as main model.*

Figure 11: Marginal effects without India



fit with the theoretical model, the effect sizes and effect certainties of the differentiation coefficient and the interaction decrease, which further corroborates the model built. In other words, if the results would have remained the same or even grown stronger when omitting India, despite it being a case that is illustrative of my theory, the empirical model may not reflect the theoretical model very well and instead pick up unrelated patterns. This slight decrease in the effect strength is therefore promising, given that the remaining sample still exhibits the expected pattern with a high probability. For a visualisation of the uncertainty, see Figure 11. Finally, Model XII does not impute any missing values in the dependent variable to signal that the results are not driven by “artificial” data.

Table 9 on page 71 tests additional empirical implications of my theoretical model. I expect a regime’s intention to coup-proof to moderate the effect of the differentiation index: a higher intention to coup-proof may signal that the increased fragmentation that the index picks up could stem from counterbalancing rather than specialisation, and vice versa. Although I adjusted the index to rather capture specialisation than counterbalancing, it still measures some form of fragmentation, and therefore

Table 8: Alternative specifications III

	Dependent variable:			
	Security Defection			
	IX: *more controls*	X: *quadr. term*	XI: *India omitted*	XII: *DV not imputed*
Differentiation	−0.570	1.092	−0.331	−0.590
	(3.49%)	(72.48%)	(22.48%)	(1.79%)
Coup Risk	−1.071	−0.656	−0.887	−1.107
	(6.73%)	(22.28%)	(15.41%)	(5.07%)
Diff∗Coup Risk	0.209	0.121	0.178	0.230
	(89.87%)	(72.42%)	(80.71%)	(92.84%)
Diff∗Diff		−0.172		
		(17.01%)		
GDP/c	−0.628	−0.642	−0.667	−0.612
	(4.63%)	(2.99%)	(2.64%)	(3.26%)
Polity2	−0.066	−0.065	−0.068	−0.067
	(2.52%)	(2.29%)	(1.72%)	(1.67%)
Military Exp.	−1.244	−0.995	−0.976	−0.978
	(0.54%)	(0.87%)	(1.10%)	(0.88%)
Violent Campaign	0.415	1.189	1.223	1.183
	(79.27%)	(99.73%)	(99.77%)	(99.78%)
Regime Support	0.033	0.187	0.074	0.201
	(54.10%)	(68.33%)	(57.10%)	(70.38%)
Campaign Support	1.305	1.454	1.463	1.437
	(99.79%)	(99.99%)	(99.98%)	(99.98%)
Population Size	−0.175			
	(26.97%)			
Ethnic Diversity	0.356			
	(76.72%)			
Religious Diversity	1.790			
	(99.99%)			
Intercept	*7.381*	*2.273*	*5.331*	*5.974*
	(1.85)	(1.62)	(1.67)	(1.75)
Observations	928	928	892	592
Note:	*Posterior means shown. Percent positive in brackets. Simulation & intercept specifications same as main model.*			

remnants of coup-proofing have to be expected. However, the coup risk index by Belkin and Schofer (2003) that I use to capture the regime's intention to coup-proof came under increased scrutiny in recent

years, especially for arbitrarily weighing its constituent parts equally (cf. Sudduth 2017). While this measurement error ought to produce more conservative estimates and may increase confidence in the existing findings, additional "sanity checks" are desirable. Therefore, I substitute the coup risk measure with alternative operationalisations that may also be able to moderate the index' effect based on the regime's desire to coup-proof.

First, democracies are less inclined to engage in coup-proofing than autocracies or anocracies (Pilster and Böhmelt 2012).[46] The main reason is the high legitimacy of the civil government, and strong civil society, both of which decrease the incentive to seize power and increase governance costs for a military that brought about an irregular change in leadership (Pilster and Böhmelt 2012; Staniland 2008).

Model XIII interacts the differentiation index with the Polity2 score. The directions of the coefficients are as expected: increased Polity2 values (more democratic) lead to a stronger negative effect of differentiation, while negative Polity2 scores contribute to an attenuation of the effect (more counterbalancing). However, drawing inference from this model would be imprudent, since the interaction coefficient is very small and experiences large uncertainty. This could be due to low face validity, or effect heterogeneity across different values of Polity2. The current model assumes the steps from autocracy through anocracy to democracy to have a linear effect on the coefficient of differentiation (in logged odds). It is perceivable that only when a certain level of stability is reached, i.e. "established" democracies, levels of coup-proofing intention start to systematically decrease. Therefore, as illustrated by Pilster and Böhmelt (2012), the Polity2 measure is re-coded into a dummy in Model XIV. Polity2 scores of 6 and higher (stable democracies) equal 1, while values below 6 (anocracies and autocracies) are coded as 0. This cutoff is arbitrary, and drawing it at neighbouring values does not change the results. Comparing the marginal effect of differentiation in democracies (-0.934) with less democratic states (-0.117), the difference is sizeable. Finally, new work by Escribà-Folch, Böhmelt and Pilster (2019) suggests that personalist autocrats are especially inclined to engage in coup-proofing. Taking advantage of these findings, Model XV uses a dummy for personalist autocracy (Geddes, Wright and Frantz 2014) for the interaction whilst controlling for Polity2 score. Again, the effect's direction is as expected, and its size and uncertainty very promising.

46. This is not to deny considerable variation in civil-military relations within regime types (Brooks 2008, 2006). On average however, Pilster and Böhmelt (2012) find substantive differences between the groups.

Table 9: Alternative specifications IV

	Dependent variable:		
	Security Defection		
	XIII: *Polity2 interaction*	XIV: *Binary interaction*	XV: *Personalist interaction*
Differentiation	−0.490	−0.117	−0.719
	(4.71%)	(37.36%)	(0.96%)
Polity2	0.010		−0.032
	(52.22%)		(18.46%)
Diff∗Polity2	−0.017		
	(32.73%)		
Democracy		2.629	
		(87.30%)	
Diff∗Democracy		−0.817	
		(6.00%)	
Personal Aut.			−3.137
			(14.89%)
Diff∗Pers. Aut.			0.987
			(91.93%)
GDP/c	−0.523	−0.506	−0.430
	(5.59%)	(6.18%)	(9.06%)
Military Exp.	−0.955	−0.970	−0.840
	(1.00%)	(1.22%)	(2.46%)
Violent Campaign	1.219	1.199	1.126
	(99.79%)	(99.67%)	(99.56%)
Regime Support	0.100	0.139	0.197
	(60.52%)	(64.37%)	(70.43%)
Campaign Support	1.427	1.463	1.426
	(99.99%)	(99.98%)	(99.98%)
Intercept	*4.780*	*3.383*	*4.655*
	(1.59)	*(1.58)*	*(1.50)*
Observations	928	928	928
Note:	*Posterior means shown. Percent positive in brackets. Simulation & intercept specifications same as main model.*		

4 Force structure and local peacekeeping effectiveness: Micro-level evidence from sub-Saharan Africa

Co-authored with Dr. Deniz Cil, post-doctoral researcher at the University of Maryland

In recent years, researchers have shifted their focus to studying the effects of peacekeeping in a geographically and temporally disaggregated manner. One of the factors that is yet to be fully examined is the variation among peacekeeping troops at the local level, and its impact on peacekeeping effectiveness. Specifically, peacekeeping troops greatly vary across two dimensions: unit types, e.g. infantry, engineering, aviation etc., and their country of origin. We argue that mixing different unit types greatly increase peacekeepers' specialization in skills and equipment, improving peacekeeping effectiveness in preventing battle-related violence. However, we also argue that this effect is strongly moderated by cultural diversity among troop contributing countries (TCCs), i.e. linguistic distance, which exacerbates coordination problems among troops. We test the propositions derived from our theoretical discussion by drawing on new subnational peacekeeping data, Geo-PKO (Cil et al. 2020). Our analysis on UN peacekeeping bases across Africa from 1994 to 2014 shows that force structure at the local level significantly impacts peacekeeping effectiveness.

Keywords: peacekeeping, effectiveness, composition, diversity, subnational

Article submitted for consideration to a peer-reviewed journal.

4.1 Introduction

The United Nations (UN) peacekeeping missions have become more multidimensional, undertaking new and complex tasks in active conflict situations that include directly engaging armed actors to neutralize them (e.g. MONUSCO), demobilizing and disarming armed groups, deterring or stopping hostilities between combatants in addition to their primary task of protecting civilians.These task require an unprecedented diversity in skills and equipment. Yet, peacekeeping bases vary greatly in their troop composition: some only have infantry troops, while other bases include engineering contingents, signals corps, maintenance, special forces, aviation, or other functionally distinct units, all of which can undertake different tasks based on their specialization. In addition to being composed of increasingly diverse units, UN peacekeeping missions are also becoming increasingly diverse culturally. As of October 2018, 124 nations were contributing to 14 active UN Missions around the globe. Yet, some bases include troops from only one troop contributing country (TCC) while others can house troops from as many as nine different countries. In this paper, we ask "how do these two dimensions of UN troop diversity, *unit diversity* and *TCC diversity*, impact local peacekeeping effectiveness to prevent battlefield violence?"

We argue that the presence of troops from a diverse set of unit types enable division of labor in the area of operation, and increase peacekeeping effectiveness to prevent battle-related violence, compared to a base composed of troops from only one unit type. Specifically, we argue that if troops from different units, e.g., aviation, engineering, or signals, are present, they can increase local peacekeeping capacity by performing tasks they are specialized at, e.g., conducting aerial surveillance, improving roads for patrolling, or establishing communications with other peacekeeping units. However, in some bases, each distinct unit is deployed by a different TCC (e.g. Pakistani infantry, Brazilian engineers, Guatemalan aviation), while in others a single TCC provides all unit types (e.g. Brazilian infantry, engineers, and aviation). We argue that coordination among units within an area of operation is easier when they come from the same TCC, or culturally similar TCCs: having troops from multiple TCCs, and especially if they are linguistically very diverse, induces communication problems and increases the chance of incompatible training and equipment. Therefore, we expect the beneficial effect of having different unit types to be moderated by the diversity of TCCs in a given location. Meanwhile, having diverse units that are able to effectively coordinate their actions is likely to improve peacekeepers' ability to address and

deter violence among armed actors by enabling them to take swift and decisive action, as well as hold and monitor larger buffer zones.

We test our propositions using a new dataset, Geo-PKO (Cil 2019), that records subnational troop deployments for all UN missions in Africa between 1994-2014. Employing Covariate Balancing Generlized Propensity Scores (CBGPS; Fong, Hazlett and Imai 2018) across different model specifications, we find support for our argument. Unit diversity, i.e., number of functionally distinct units, improves local peacekeeping effectiveness in reducing the likelihood of nearby combat events. This effect, however, is moderated by TCC diversity. If the different units types are also culturally distinct, effectiveness does not improve.

We also provide supporting evidence from semi-structured interviews with a small but diverse sample of former and current peacekeepers. These interviews highlight the benefits of unit diversity, as well as the challenges induced through TCC diversity, from the perspective of peacekeepers. They allow us to unpack how the daily peacekeeping tasks are improved by unit diversity as well as the ways in which cultural differences impact peacekeepers' ability to work together.

This paper makes several contributions to existing scholarship. First, it puts forward a theory about how troop diversity impacts conflict processes. Previous studies examine how broad types of peacekeepers, such as troop, police, or observer units, perform different tasks and therefore are better able to fulfill different peacekeeping mandates. We extend this literature by offering a more nuanced theory of peacekeeping effectiveness from the perspective of military cooperation and coordination among contingents from different countries, and go beyond the assumption that all peacekeeping troops are equal. To our knowledge, this paper is the first to disaggregate peacekeeping troops by unit type and systematically evaluate the effect of their task specialization on peacekeeping performance. However, we are aware that this specialization often comes with a high diversity in TCCs. Therefore disentangling these two dimensions, and focusing on the interaction between them, is paramount. With this approach we offer a more complete picture of unit composition.

Second, we build on the emerging research on the subnational analysis of peacekeeping effectiveness. With the exception of a few recent quantitative studies (Ruggeri, Dorussen and Gizelis 2018; Fjelde, Hultman and Nilsson 2019; Cil 2019), and several disaggregated case studies (Costalli 2014; Autesserre 2014), the study of peacekeeping effectiveness at the subnational level is still very limited.

This is especially important, because even if at the national level some missions may seem similar, at the local level there is great variation in troops' geographic concentration, their composition, and their capabilities, which may obscure or confound findings.

Lastly, and most importantly, our theory and results offer insight into how peacekeeping works locally, and have policy implications for the subnational deployment patterns of future UN peacekeeping missions. Given that the force commanders have to make various trade-offs in their deployment decisions, it is crucial to maximize the potential of each unit in peacekeeping missions. Our results show that while additional specialized units significantly improve effectiveness, units' cultural distance from each other is a strong moderating factor that can offset the positive impact of unit diversity.

4.2 Local peacekeeping effectiveness

In the last two decades, research re-evaluated peacekeeping effectiveness by utilizing quantitative methods and taking into account the non-random nature of peacekeeping deployments. Studies have shown that peacekeeping missions increase the duration of peace following civil wars (Doyle and Sambanis 2006; Fortna 2008), even though they are deployed in more severe conflicts with higher risk of recurrence (Fortna 2008; Gilligan and Stedman 2003; Gilligan and Sergenti 2008).

Recent research, with the availability of new data (Kathman 2013), started to examine peacekeeping effectiveness in a disaggregated manner, going beyond the presence and type of mission in general. Instead, it focused on personnel types, i.e. troops, police, and observers, and mission capability, i.e., number of peacekeepers. For example, higher numbers of peacekeeping troops and police units lead to lower numbers of battle-related and civilian deaths in ongoing conflicts (Hultman, J. Kathman and Shannon 2013). Larger troop presence in post-conflict countries is also associated with longer duration of peace and less civilian victimization (Kathman and Wood 2014; Hultman, J. D. Kathman and Shannon 2016). Others examine how missions' TCC diversity, as well as the distance between the TCCs and the local population in terms of geography, language, and culture, impact peacekeeping effectiveness at the mission level (Bove and Ruggeri 2016, 2019).

While the extant research on peacekeeping offers valuable insights into how peacekeeping missions operate and when they are effective, several important topics remain unexplored. One of the main drawbacks of this literature, with the exception of the few studies we discuss below, is the cross-mission

and cross-country analysis of peacekeeping effectiveness. This is problematic, since both peacekeeping deployments and conflict processes vary greatly at the subnational level. Findings at the aggregate national and mission level may be over- or underestimating the actual effect of peacekeeping on the ground. As a result, in the last several years, the focus of the literature has shifted towards a more geographically disaggregated study of peacekeeping, offering better overlap between the theoretical mechanisms and their quantitative assessment.

For example, Ruggeri, Dorussen and Gizelis (2017) show that the deployment of peacekeepers in larger numbers – around 500 – reduce the probability that the conflict will continue at the local level. This finding provides initial evidence that the previous studies at national level do not overestimate peacekeeping effectiveness. Similarly, Fjelde, Hultman and Nilsson (2019) show that larger numbers of peacekeepers lower the likelihood of civilian targeting at the local level, especially by rebel forces. Both studies also show that peacekeepers tend to be deployed in active conflict regions within a country, in line with previous studies that point to similar deployment patterns at the national level.

Additionally, existing research shows that within-mission TCC diversity is an important factor that improves mission effectiveness in reducing one-sided violence (Bove and Ruggeri 2016). On the other hand, at the mission level, larger religious and linguistic distance between the peacekeepers and the locals leads to higher levels of one-sided violence and battle-related deaths (Bove and Ruggeri 2019). Given the variation in how peacekeepers are distributed within a country, however, it is important to examine whether similar patterns hold at the local level. With the exception of Cil (2019) that examines unit composition at the local level and its effect on both battle-related deaths and one-sided violence (OSV), there are no other quantitative studies that examine unit composition at the subnational level across multiple missions. This is a significant gap in the current literature for several reasons.

First, at the mission level, TCC diversity contributes to a decrease in violence levels. This may be either due to specialization, and different skill sets and technologies complementing each other and improve overall success, or due to increased diplomatic leverage through a higher number of external monitoring actors. Therefore, whether and how diversity influences effectiveness at the local level remains to be examined. Specifically, deploying troops from a diverse set of countries can also induce coordination problems, which may be more observable at the local level. Second, past emphasis on troop's cultural diversity overlooked an equally salient factor determining local variation of peacekeeper's

composition: functional units. Across missions, TCCs vary in which kind of functional units they are set to contribute. Therefore, which training and equipment peacekeepers receive from their home country is not only dependent on the country itself, but for the most part on the kind of task for which task they are being deployed within each mission. Consequently, to determine the effect of troop composition on effectiveness, investigating the effect of unit diversity is at least as relevant as evaluating the effect of TCC diversity. For example, whether a peacekeeping base only holds infantry troops, or also engineers and special forces, can make a big difference in their local readiness and effectiveness. The potential effects of this variation is impossible to observe at the aggregate level.

To our knowledge, this paper is the first one to systematically disentangle, both theoretically and empirically, the beneficial effect of peacekeepers' functional differentiation from the potential coordination problems induced by cultural diversity. To this end, we examine the effect of troop diversity on peacekeeping effectiveness along two dimensions: unit diversity and TCC diversity. Specifically, presence of different units such as infantry, engineering, transport, or aviation can significantly improve peacekeeping performance. Yet, this effect may be moderated by the extend to which these units are able to cooperate and coordinate across languages and cultures.

4.3 Force structure in peacekeeping

As discussed above, force structure is a relatively new topic in peacekeeping literature. Until the seminal study by Bove and Ruggeri (2016), the relevance of peacekeepers' composition in terms of their backgrounds, skills, and equipment, was mostly overlooked by the quantitative literature. In the wider context of military studies however, the role of force composition in military effectiveness is long established (starting with Huntington (1957)). Specifically, force employment, i.e., how the different units such as ground and air forces are used, is critical for achieving favorable battlefield outcomes in both interstate as well as internal conflicts (Biddle 2004; Caverley and Sechser 2017). Given the importance of how forces are utilized in the battlefield for successful outcomes, it is crucial to examine if and how force composition impacts peacekeeping effectiveness, as UN peacekeeping operations have become more complex and multidimensional.

While these robust peacekeeping operations have protection of civilians and use of force for self-defense as part of their mandate, their effect on violence go beyond protection of civilians. Specifically,

peacekeepers deter or reduce violence between two armed actors through two main mechanisms, i.e., mitigating commitment problems and increasing the costs of continued fighting (Hultman, J. Kathman and Shannon 2014). They accomplish this by undertaking increasingly more "military" tasks such as capturing and holding checkpoints, separating and disarming combatants, patrolling, and intelligence gathering. Peacekeepers are also deployed in active conflict countries, a divergence from the early missions that were largely composed of military observers deployed to monitor existing ceasefire treaties.

Further peacekeepers usually carry out these tasks under conditions of power asymmetry requiring local intelligence (Howard and Dayal 2018; Hultman, J. Kathman and Shannon 2014, 2013; Fortna 2008). On the ground, this means that peacekeepers undertake tasks ranging from policing, mobile and site protection, humanitarian aid, local trust building, to targeted interventions. In other words, peacekeepers engage in conflicts with similar dynamics as the ones counterinsurgents face, and undertake similar tasks at times (Pilster, Böhmelt and Tago 2016; Friis 2010; Celeski 2009; Lyall 2009; Kalyvas 2006; Wickham-Crowley 1992; Galula 1964), with the apparent exception of offensive use of force (Howard 2019; Howard and Dayal 2018). As a result, insights from military studies on force composition can be helpful to understand how it may impact peacekeeping effectiveness.

The range of peacekeeping tasks suggests that the implementation of robust peacekeeping mandates requires a flexible set of skills and equipment, similar to counterinsurgency operations. The effectiveness of counterinsurgency operations does not depend on a particular toolbox of tactics and strategies that should always be employed. Rather, it depends on the actors' ability to select the right procedures, accounting for the specific characteristics of the security environment they face. To this end, forces need to rely on local intelligence, and have the training and equipment that is specialized in the distinct security task they need to accomplish. As a result, a single conventional force from one specific training background is unlikely to cover all bases (Dworschak 2020; Dworschak 2019; Pilster, Böhmelt and Tago 2016; Lyall 2009; Kalyvas 2006; Galula 1964, 66). The same applies to peacekeeping. UN troops reduce battlefield violence by separating combatants, monitoring frontlines, dismantling checkpoints, and creating and holding buffer zones (Hultman, J. Kathman and Shannon 2014). Maintaining multiple parallel units can increase the chance of having a diverse set of skills and equipment in one place (Bove and Ruggeri 2016).

The presence of specialized units such as helicopter support or engineers is likely to improve troops' ability to patrol buffer zones, gather information on troop movements, or fortify observation posts, increasing the effectiveness to deter or reduce armed violence. When diffusing confrontation between armed actors, for example, infantry units working together with helicopter units providing close air support can more effectively reach conflict areas and more decisively halt violence. All of these tasks, however, require high levels of training, skill and coordination among units (Biddle 2004; Cil 2019). Especially in UN peacekeeping missions, units from many different countries are often deployed in the same location. Unlike many counterinsurgency operations in which troops belong to one or a few allied nations, units in UN missions have a highly diverse set of languages and cultures, which may pose a challenge to efficient coordination among troops.

As a result, we argue that the local troop composition in UN missions should be considered as a two-fold concept, encompassing both *unit diversity* and *TCC diversity*. We argue that these two dimensions of troop diversity are likely to affect peacekeepers' effectiveness in preventing and deterring battle-related violence. First, a diverse set of functional units improves effectiveness at the local level by increasing troop specialization, and enabling a division of labor. This prevents or reduces violent incidents between armed groups by allowing peacekeepers to react promptly, demonstrate superior capacity, and hold buffer zones more effectively. Second, we argue that TCC diversity moderates this relationship by inducing coordination problems among troops due to communication problems and cultural incompatibilities. Specifically, high levels of training and skills is required for effective coordination among different units to make use of the full potential of each unit. TCC diversity can hamper these efforts and reduce any positive impact of unit diversity on deterring battlefield violence. In the following section, we elaborate further on how unit and TCC diversity can impact daily peacekeeping operations, and the overall effectiveness to reduce armed violence.

4.3.1 The need for cooperation and coordination

One of the mechanisms through which troop diversity impacts peacekeeping effectiveness positively is the division of labor among different functional units. When there are multiple distinct units in a location, each brings different strengths to the table. Infantry units are normally the "baseline", conducting patrols and offering force protection. Deployment of engineering units is also common, as they have the training

and machinery to maintain the base, build and repair roads and bridges, drill wells, etc. Examples of other unit types include reconnaissance units which collect local intelligence to inform operations, and demining units which clear roads of mines and IEDs.

Thus, a peacekeeping base benefits from dividing various peacekeeping tasks based on each unit's training background, equipment, and experience. If the infantry unit takes care of establishing operational security, engineers and maintenance units are free to focus on repairing vehicles and infrastructure, or fortifying positions. Through this division of labor, each unit's strength can be fully exploited and the overall effectiveness of peacekeepers in a given base can be improved. A prime example of the direct effect of diverse unit types on daily operations is the presence of aviation units. Helicopter units can increase the ability to do air patrol to determine troop movements and provide close air support for ground troops in case of confrontation. Take Location A where two units, two infantry companies, are deployed, and compare it to Location B where two units, one infantry company and one aviation squad, are deployed. Even though the total number of troops is higher in Location A,[47] we expect that the effectiveness of peacekeeping troops to be much higher in Location B, as they have increased capacity to monitor the vicinity, collect intelligence, and provide logistics and air support using the aviation unit.

Presence of additional specialized units that can focus on important support tasks is likely to boost the effectiveness of troops in reducing battle-related violence by allowing infantry troops to work more efficiently, i.e., patrol larger areas faster. The presence of infantry troops, in turn, can provide the necessary force protection for supporting units. For example, an engineering unit repairing a road must have other peacekeepers guarding them during their labor. While each unit is required to be "self-sufficient", i.e. to be able to provide their own force protection, this would mean assigning a trained engineer for protection, which constitutes an inefficient use of human capital. When these different units work together, they can accomplish tasks more effectively and operate smoothly.

As soon as coordination and cooperation problems occur, however, they have to allocate existing manpower to cover tasks outside of their expertise. One of our interviewees, a combat engineer, recalls an incident in which the infantry battalion had to leave the current area of operation, and lamented that "[..] [we] had to put some extra engineer guys form our platoon to work as infantry guys, because we could not rely on the [country]'s battalion."[48]

47. Two companies would be composed of approximately 300 troops, while one company and a squad would total to 160-175 troops.

48. The incident came about due to a cooperation problem with the troops from that specific TCC.

This example brings us to the second dimension of diversity, i.e., TCC diversity. While division of labor between specialized units can increase effectiveness, for this to work troops need to be able to seamlessly coordinate their actions. We argue that TCC diversity of UN troops may render them less able to coordinate their actions efficiently, impairing the effect of unit diversity. Specifically, we argue that there are three main obstacles to effective coordination among different specialized units who are from different countries.These are, (1) linguistic differences, (2) cultural differences, and (3) differences in military training and standard operating procedures (SOP). When the troops do not speak the same language, and have to use translators, for example, effective and timely communication between units can be hampered. When troops come from different cultural and religious backgrounds, they may have different methods of approaching certain tasks. Further, these differences may cause tension among troops, preventing effective division of labor. Lastly, although troops may have similar cultural background, they may have differences in their military training and operating procedures that can make it difficult for two units to work together.

Linguistic differences lead to communication problems, imposing obvious obstacles to daily operations. Units from different TCCs that need to coordinate their actions require cross-TCC compatible radios, which is not always ensured. In addition, they may need a translator, which is likely to increase the time to complete certain tasks and reduce overall effectiveness of the operation. Further, a translator may not always be available, making it nearly impossible to coordinate.

Beyond these communication barriers, cultural and attitudinal differences (perceived or actual) in the conduct of military operations, the interpretation of the peacekeeping mandate, and the use of force, can hinder cooperation or worse, deter future attempts to cooperate within a base. Reducing battle-related deaths through separation of combatants requires each unit to coordinate its actions. Carrying out complex operations to secure a location and neutralize armed actors, requires coordination and cooperation among infantry troops. To effectively carry out both defensive and offensive operations, however, high levels of military training and flexibility in decision making (for local commanders) is often necessary (Biddle 2004). Unlike national militaries with clear rules of engagement, command and control structures, units in a UN peacekeeping mission are supplied by different TCCs. Each TCC joins the mission with a different background in training that can vary based on national doctrine and combat experience. As a result, each TCC can interpret and carry out mandates in somewhat different

ways (Soeters, Poponete and Page Jr. 2006; Soeters 2018; Ruffa 2014). As the evidence from operations of other multinational coalitions, such as NATO's International Security Assistance Force (ISAF) show, coordination and agreement amongst soldiers about the appropriate protocols to follow is crucial (Auerswald and Saideman 2014). Similar problems of coordination that hinder daily operations of peacekeepers are also recorded across many UN missions (Bove and Ruggeri 2016).

This discussion leads us to our hypothesis on the beneficial effect of unit diversity, conditional on coordination impediments induced by TCC diversity:

H_A: Unit diversity reduces the likelihood of battle-related violence,
only if peacekeepers experience low coordination problems.

4.4 Research design and empirical evidence

4.4.1 Variables and Data

In order to test the hypotheses presented above, we use a new subnational peacekeeping deployment datatset. The Geocoded Peacekeeping Operations (Geo-PKO) data records disaggregated information on individual peacekeeping bases, including the types of units (infantry, engineers, signals corps, etc.) and troop contributing countries (TCCs) present in each location (Cil 2019). This dataset covers all UN missions deployed to African countries between 1994 and 2014, with a total of 27 missions. The coding of the Geo-PKO dataset is based on UN deployment maps. The data includes 417 maps covering 125 mission-years and 9446 unique mission-locations. The deployment updates differ from mission to mission, and not all missions have monthly deployment updates. The mean number of maps per mission-year is three. 17% of mission-years have one deployment map, 15% have two, 22% three, and 46% have four or more maps (maximum of eight) in a given year. In line with the previous studies using an earlier version of this dataset (Fjelde, Hultman and Nilsson 2019), we assume that there is no change in the deployment pattern of the peacekeepers until a new map is available.[49]

In addition, to capture the coordination problems conceptualized above, we draw on data that denotes the linguistic distances between troops (Spolaore and Wacziarg 2016). Linguistic distances are calculated based on tree-based methods, capturing how phonetically and historically "close" languages

49. Missingness is limited to very few observations in which TCCs are "unknown". For the few cases in which a TCC's troop size is not recorded, we assume a minimum contribution of ten soldiers. These choices do not systematically influence our results.

are. Bove and Ruggeri (2019) used the same data, however they used it for measuring the average distance between peacekeepers and the local population. Because we are examining the coordination problems between peacekeepers, we use the data to measure average linguistic distances between all TCCs operating in the same location. Finally, we use the Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset (GED) to measure our dependent variable, the likelihood of a combat event between conflict actors to occur in a given location (Sundberg and Melander 2013).

As our unit of analysis, we aggregate both datasets to grid-month level with a spatial resolution of 0.5x0.5 decimal degrees (Tollefsen, Strand and Buhaug 2012). We restrict our sample to only those grid-month observations that occur in a country and year of active conflict, as recorded by UCDP. In addition, since we are interested in comparing the effectiveness of different peacekeeping deployment features, we only consider those grid-month observations in which peacekeepers were actually present. In other words, our analysis compares the likelihood of battle events between deployment locations with large troop diversity to those with little or no troop diversity, and leaves out locations where there is no UN presence (i.e., where no treatment can occur). Figure 12 on the following page provides a spatial overview of our grid cell sample.

For estimation we use logistic regressions, since our dependent variable is binary. We report all results with robust standard errors clustered on grid level.[50]

Dependent variable

Using the UCDP GED dataset, we identify grids with and without battle-related deaths in a given month in all conflict countries and years that crossed UCDP's total 25 battle-related death threshold. The dependent variable is therefore coded 1 when there is at least one battle-related death observed in a given grid-month. 123 observations (3% of our sample) are coded 1.[51]

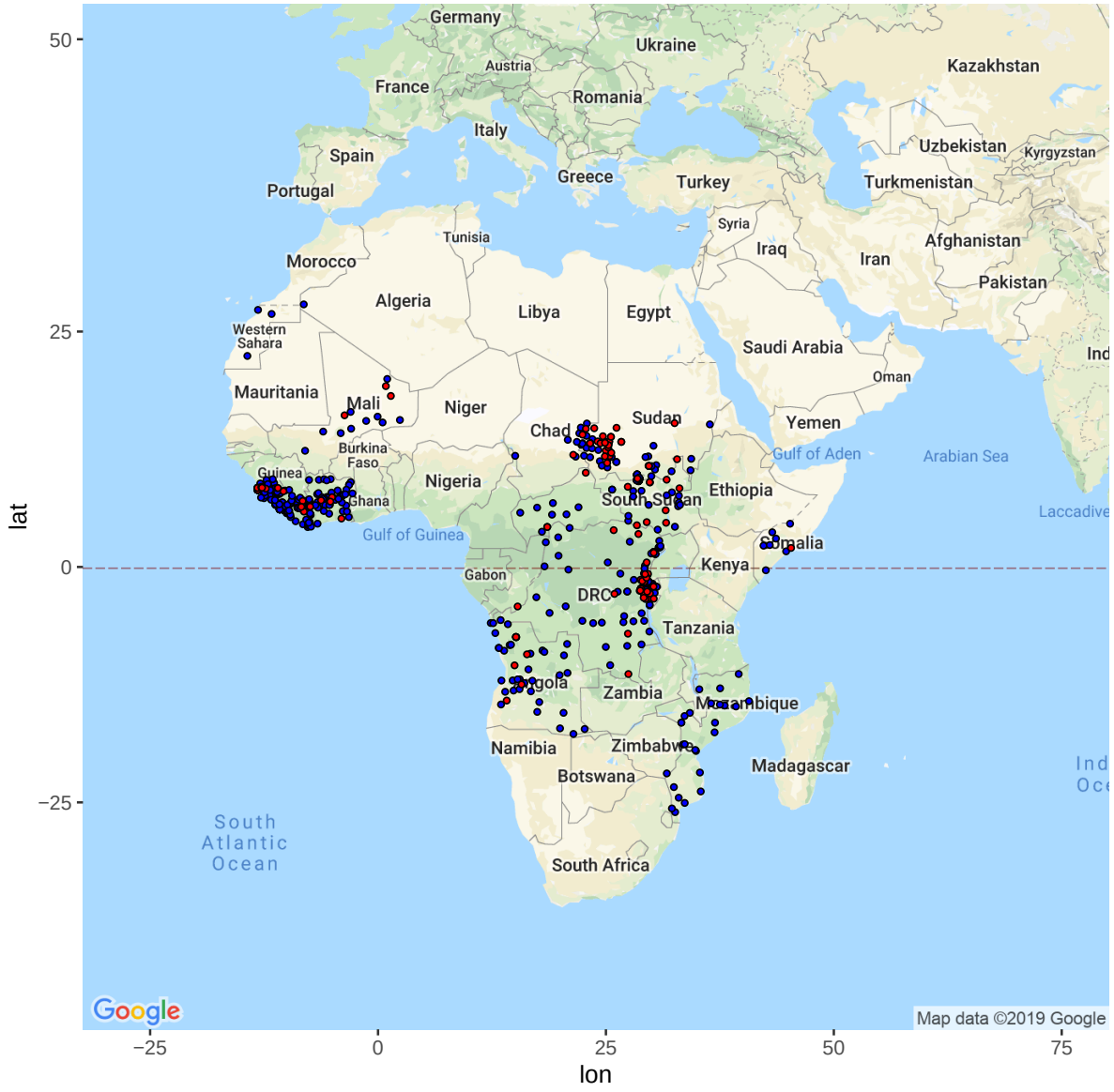
Independent variables

We use the Geo-PKO dataset and distance data to measure our independent variables. Following our theory and hypotheses, there are two types of diversity that interact with each other.

50. These are the most conservative standard errors. Clustering on the more aggregate mission level, or not clustering at all, both lead to less conservative estimates. Also, re-estimation using a Linear Probability Model yields similar results.

51. We rerun our models using Firth Logit that is appropriate for rare event data. The results, which are listed in the Online Appendix, remain unchanged. In addition, results remain robust to an alternative threshold of five deaths for measuring occurrence of battle-deaths (Fjelde, Hultman and Nilsson 2019).

Figure 12: Mapping our sample of PKO grid cells, 1994-2014

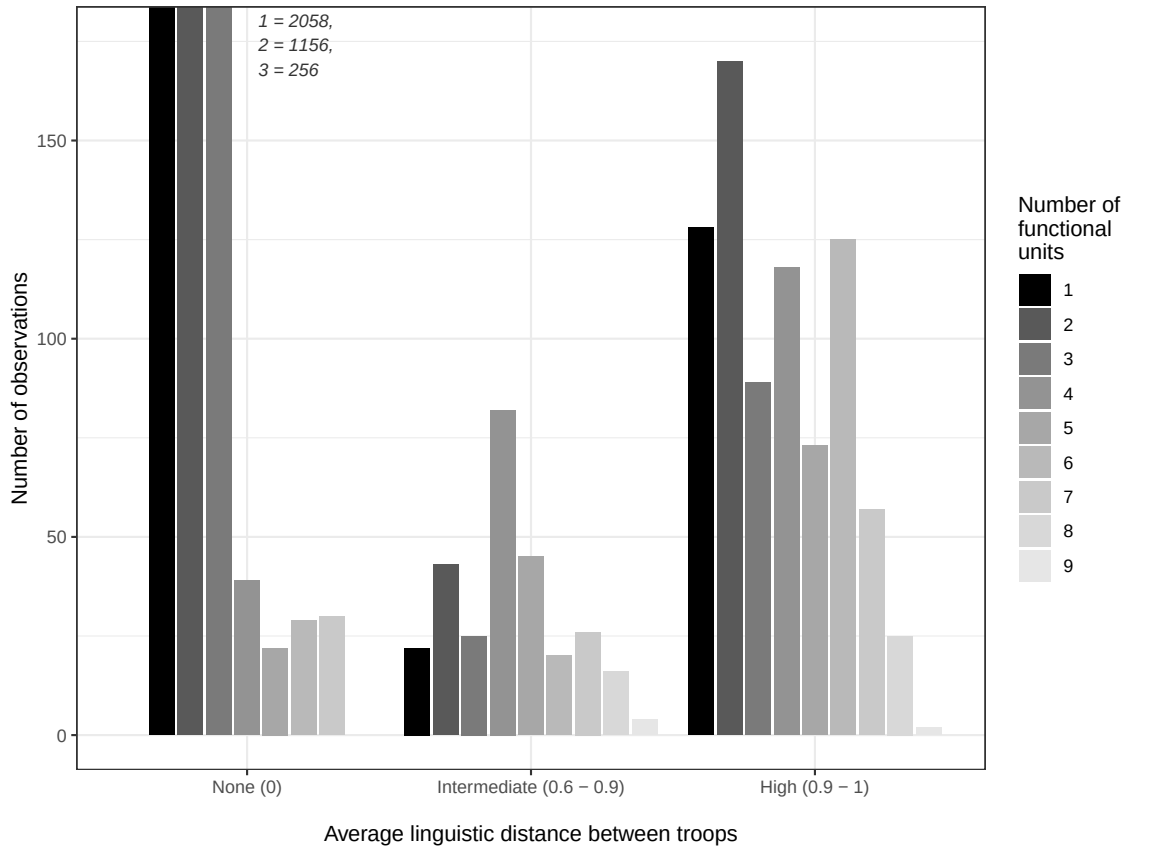


Red: PKO grid cell that did experience conflict at some point in time, as coded by UCDP GED.
Blue: PKO grid cell that did not experience conflict.

First, there is unit diversity, which implies units' specialization on distinct tasks and a division of labor. In other words, it is the number of unique unit types in a given location, each fulfilling a different function. Possible unit types deployed in PKO bases are infantry, engineers, signals corps, transportation, riverine units, helicopter support, special forces, medical units, maintenance, reconnaissance, aviation, military police, demining units, (civilian) police units, and military observers (Cil 2019). To code unit diversity, we consider the number of unique units in any given grid-month observation. For example, if a base has a Brazilian infantry unit, a Brazilian engineers unit, and an Ethiopian engineers unit, unit diversity in this case means that the troops on that base are specialized as both infantry and engineers.

Therefore it would be coded as 2. 47% of observations in our sample contain one unique unit type, 29% have two unit types, 8% with three, 5% with four, 3% with five unique units, and the remaining 8% with six to nine (maximum) distinct unit types.[52]

Figure 13: Unit diversity and linguistic diversity



"Observations" are PKO grid cells per month. The first three bars (observations with troops from one, two, and three functional units with no linguistic distance between each other) are clipped off to keep the y-axis to a reasonable scale. The numbers of observations are shown at the top of the graph.

Second, there is TCC diversity of troops, inducing coordination problems between units speaking different languages, and coming from different military and cultural backgrounds. We operationalize coordination problems in two complementary ways, expecting homogeneous effect directions across the different measures. First, we use the average linguistic distances between peacekeepers.[53] Linguistic distances are coded as the mean distances between TCCs in a specific grid-month observation. Each TCC has a unique linguistic distance from every other TCC present in that location. We take the mean

52. It appears plausible that individual unit types may contribute differently to effectiveness. While disaggregating our measure even further and exploring this potential effect heterogeneity is not the focus of this paper and goes beyond its scope, it is left for future research to examine.

53. As a robustness check, we also use religious distance instead of linguistic distance, expecting a similar effect. These supplementary analyses are included in the online appendix and provide additional corroboration.

distance for each TCC, and then average across the TCC-specific means within each grid-month observation. In other words, we capture the grand mean of all linguistic distances between TCCs in a given location and month. The distances range between 0 (no distance) and 1 (maximal distance). Second, as a stress test, we substitute the distance measure with a conservative binary indicator of whether troops in a given location are all from the same TCC, or whether they are from different TCCs. This captures whether peacekeepers on a base have to cooperate across different TCCs (proxying for potential cultural differences), or whether all peacekeepers come from the same country.[54] 77% of our sample exhibits no linguistic distance, where units are provided by only one troop contributing country. The remaining 23% consist of varying degrees of distances, ranging up to 1, which denotes deployments of a very diverse set of TCCs in a location. See Figure 13 on the previous page and Figure 14 on the following page for a visualization of these distributions.

A potential issue transpiring from these figures is that there are no observations with eight or nine functional units from the same TCC. To make sure that our results are not driven by these leverage points and lack of common support, we rerun our analyses without these observations. Results for these subsample analyses, which can be found in the Online Appendix, remain similar to the results reported in the main text below.

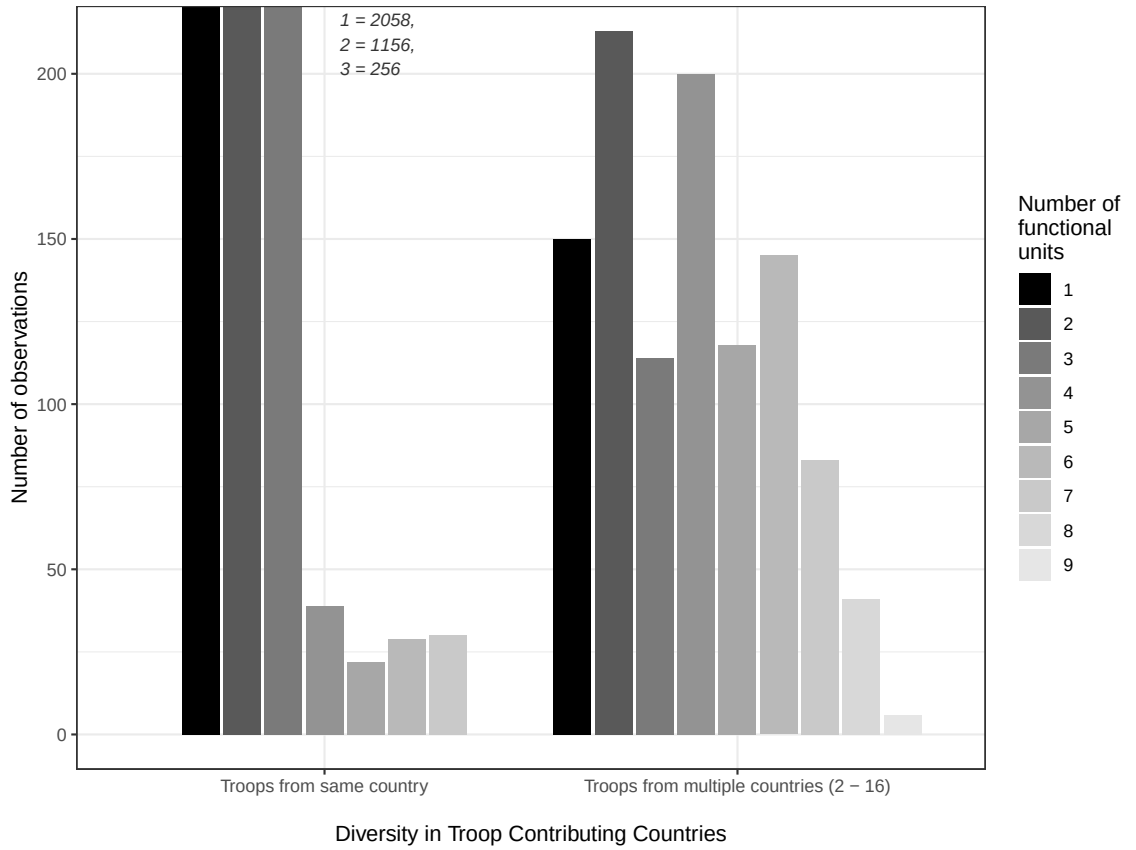
4.4.2 Threats to inference

The composition of units may be dependent on factors that also systematically influence our outcomes of interest. Previous studies show that the patterns of subnational peacekeeping deployment vary depending on the local conflict environment. Specifically, Ruggeri, Dorussen and Gizelis (2018) find that peacekeepers are deployed in areas of active conflict within a country, which is in line with the previous cross national studies that find that peacekeepers are deployed to difficult cases. They also find evidence that ease of access is another strong predictor of subnational deployment, thus offering an insight into how logistical concerns might play a role.

If the local conflict processes affect the composition of deployment, then our analysis may produce biased results. For example, the UN may deploy more diverse units in areas with more intense conflict activity. Alternatively, the UN may deploy more units from different TCCs in and around the cap-

54. An alternative operationalization in which we use the raw number of TCCs instead of the dichotomous measure yields similar results.

Figure 14: Unit diversity and country diversity



"Observations" are PKO grid cells per month. The first three bars (observations with troops from one, two, and three functional units from a single troop contributing country) are clipped off to keep the y-axis to a reasonable scale. The numbers of observations are shown at the top of the graph.

ital city where the headquarters are usually located, and where better infrastructure is present. Fewer units and TCCs may be deployed to rural locations close to borders. It may be difficult to reach these areas, or there may be fewer TCCs who are equipped to handle more complex operations undertaken in border regions.

To empirically address this issue and better estimate the effect of force composition on battle-related violence, we account for various confounding variables in our regressions. We also use weights from Covariate Balancing Generalized Propensity Score (CBGPS) matching to reduce the correlation between the confounders and our main treatment *unit diversity* while relaxing functional form assumptions across strata (Imai and Ratkovic 2014; Fong, Hazlett and Imai 2018). The CBGPS matching approach is designed to overcome the well-known dimensional limitations of traditional propensity score matching, and is suitable for continuous treatment variables. In line with previous studies, we identify several variables that capture conflict history and intensity, ease of access, and other grid-cell charac-

teristics that are likely to impact force composition as well as the underlying levels of violence between combatants.

The “treatment” is our measure of unit diversity in the previous month (t-1) in a given grid cell. Using a lagged version of unit diversity decreases concerns of reverse causality, i.e., violence levels predict how many units from different TCCs are deployed to a grid cell in the same month.[55] During the interview with a Military Chief of Staff, we specifically asked about the use of intelligence to determine *future* hot spots, and sending troops somewhere in anticipation of fighting in the near future. She/he clarified that "[The mission’s circumstances] did not leave any options to deploy troops into other areas than only those where hostilities took place at the time", and that this is probably not limited to this particular mission. Moreover, such rapid troop allocations would be largely limited to infantry contingents. Given our coding of *unique* unit types (i.e. having a single versus multiple infantries in one location does not change our measure), this would not confound our analysis.

To further address concerns of reversed causality, we include variables capturing previous conflict trends, i.e. whether or not battle events occurred in the previous three months (t-3 to t-5) in the given grid cell, as well as conflict history in neighboring grid cells (t-2) to account for spatial dependence. These variables are recorded using the UCDP GED. In addition, based on the understanding that "forward bases" in more remote and hostile areas tend to be more temporary than centrally located, long-established bases, we include cubic polynomials of time since the beginning of peacekeeping presence in a given area.[56]

We also match on the (logged) total number of troops in a grid cell. The diversity of units and TCCs is likely to be higher in locations with a large number of troops. Similarly, large numbers of troops are likely to reduce the baseline probability of battle-related violence. To account for potential spill over effects we include (logged) total number of peacekeeping troops in the neighboring cells as well. To account for the "ease of access" as explained above, we also match on the grid cell’s distance to the capital and distance to the border of the nearest land-contiguous neighboring country. In order to address other more latent factors that may impact underlying propensity of conflict and force composition, we record nightlight coverage, mountainous terrain, and whether there are excluded ethnic groups in the grid-cell.

55. All time varying variables in the matching equation are measured at t-2, thus preceding the treatment measured at t-1, to shield against post-treatment bias (Imai and Ratkovic 2014; Ho et al. 2007).

56. In a separate test we include a binary measure of whether or not an area has a headquarter. The results stay robust to this conditioning.

We use PRIO GRID to record all of these grid level variables. Lastly, we employ fixed effects on country level to capture any time invariant country- and mission-specific heterogeneity.[57]

Finally, we conduct a minor bounding exercise to learn about the sensitivity of our results in the face of any remaining endogeneity. As explained above, some areas may have received the treatment (increased number of units) due to more or less prior conflict activity. To better understand the degree to which such a reversed relationship may drive our results, we simulate 800 new datasets. In each iteration, we raise/attenuate a share of the treatment variable (assign one unit more/less) for those observations that have experienced violence in the previous three months. In other words, we assume that a share of the data we observe was shaped by reversed causality and manipulate the data to alter that share. This helps us to better understand "how much" reversed causality would be necessary to solely drive our results. The first 400 iterations adjust random samples of 20% of the potentially endogeneous observations, with 200 datasets raising the number of unit types as a result of previous violence, and 200 lowering the number. Across all simulations, there is no single iteration that turns the effects of interest substantively insignificant. Only two of all 400 interaction coefficients exhibit p-values slightly higher than 0.05. A second run with 400 iterations similarly adjusts samples of 40% of relevant observations, with the results staying robust. Here, five of all 400 interaction coefficients exhibit p-values slightly higher than 0.05. The coefficient and p-value distributions are visualized in the Online Appendix. While this does not alleviate concerns over endogeneity, it provides further corroboration that we have identified a systematic pattern.

4.4.3 Results

Table 10 on the next page shows the results from our analyses in which the dependent variable is the occurrence of battle-related deaths. There are two sets of models, each set listing results for a different moderating variable that captures potential coordination problems. Each set contains a model with and without CBGPS matching. Models 1 and 2 show a strong negative effect of unit diversity on the likelihood of battle events when there is no linguistic distance between peacekeepers. This effect is significantly attenuated when troops speak different languages in a given location (gray-shaded row). The same pattern holds for Models 3 and 4: unit diversity does not significantly contribute to

57. To further assess model dependence and address the potential of overfitting, we include models without fixed effects in the Online Appendix. Our results remain robust.

Table 10: Logistic regressions

	Occurrence of battle-related deaths			
	(1)	(2)	(3)	(4)
Unit diversity (t-1)	−0.835***	−0.868***	0.082	0.305
	(0.228)	(0.282)	(0.113)	(0.217)
TCCs' linguistic distances (t-1)	−1.803**	−2.985**		
	(0.783)	(1.179)		
Same TCC (t-1)			2.198***	2.857***
			(0.691)	(1.015)
Battle deaths(bin)(3m ave.)	1.548***	1.162***	1.470***	1.130***
	(0.230)	(0.334)	(0.229)	(0.347)
Neighb. battle deaths(bin)(t-2)	1.658***	1.602***	1.687***	1.630***
	(0.272)	(0.297)	(0.271)	(0.302)
Number of troops(ln)(t-2)	0.362**	0.470***	0.358**	0.445***
	(0.154)	(0.130)	(0.148)	(0.129)
Neighb. Troops(ln)(t-2)	−0.040	0.057	−0.068	0.052
	(0.051)	(0.068)	(0.054)	(0.070)
Mountainous Terrain	1.086	0.811	1.126	0.787
	(0.753)	(0.846)	(0.744)	(0.855)
Nightlights Coverage	5.183	3.789	7.397	4.761
	(4.716)	(7.868)	(4.767)	(8.386)
Excluded Ethnic Groups	−0.265	−0.282	−0.310	−0.317
	(0.206)	(0.250)	(0.210)	(0.254)
Distance to Capital	−0.233	−0.415	−0.137	−0.403
	(0.213)	(0.390)	(0.211)	(0.394)
Distance to Border	−0.242*	−0.228	−0.208	−0.211
	(0.130)	(0.160)	(0.129)	(0.160)
Unit diversity * Ling. dist.	1.008***	1.287***		
	(0.301)	(0.428)		
Unit diversity * Same TCC			−1.334***	−1.321***
			(0.364)	(0.452)
Constant	−2.984*	−2.214	−5.096***	−4.804**
	(1.684)	(2.771)	(1.627)	(2.440)
Observations	4660	4660	4660	4660
CBGPS matching	X	✓	X	✓
Country FE	✓	✓	✓	✓
Time cubic polynomial	✓	✓	✓	✓

Note: *p<0.1; **p<0.05; ***p<0.01

peacekeeping effectiveness when peacekeeping troops in the same location do not come from the same contributing country. In other words, the effect of unit diversity is not statistically discernible from zero when the binary variable "Same TCC" is zero. However when units are provided by a single TCC, an increased number of functional units substantially improves effectiveness (last gray-shaded row).

These effects are visualized in Figure 15 on the following page and Figure 16 on page 93. They paint a consistent picture: unimpeded unit diversity strongly reduces the likelihood of battle-related deaths in a given grid-month, from around 5% down to approximately 0%. These results underline two relevant patterns. First, functional differentiation, i.e. having multiple units specialized on distinct tasks, strongly contributes to peacekeeping effectiveness. An infantry unit which does not need to divert attention to the maintenance of the base, due to an engineering unit covering these tasks, has more resources available for patrolling, dismantling checkpoints, and force protection. However, such a division of labor requires coordination between the units. Therefore, second, the effect of functional differentiation is strongly moderated by cultural diversity, which induces coordination problems. In an emergency, aviation support can only help an infantry unit if the latter can actually call it in – that is, if language barriers are surmountable and communication equipment is compatible. In the visualizations shown here, linguistic distance is set to 0 (no distance) and 1 (maximal distance) respectively.[58]

The importance of recognizing the heterogeneity in the effect of unit diversity across different realizations of cultural diversity is further signified through the improvement in model fit when including the interaction term. Comparing model 2 with and without multiplying unit diversity and linguistic distance, the Akaike Information Criterion (AIC) improves by a value of 21 (from 1240 down to 1219) when including the interaction term.

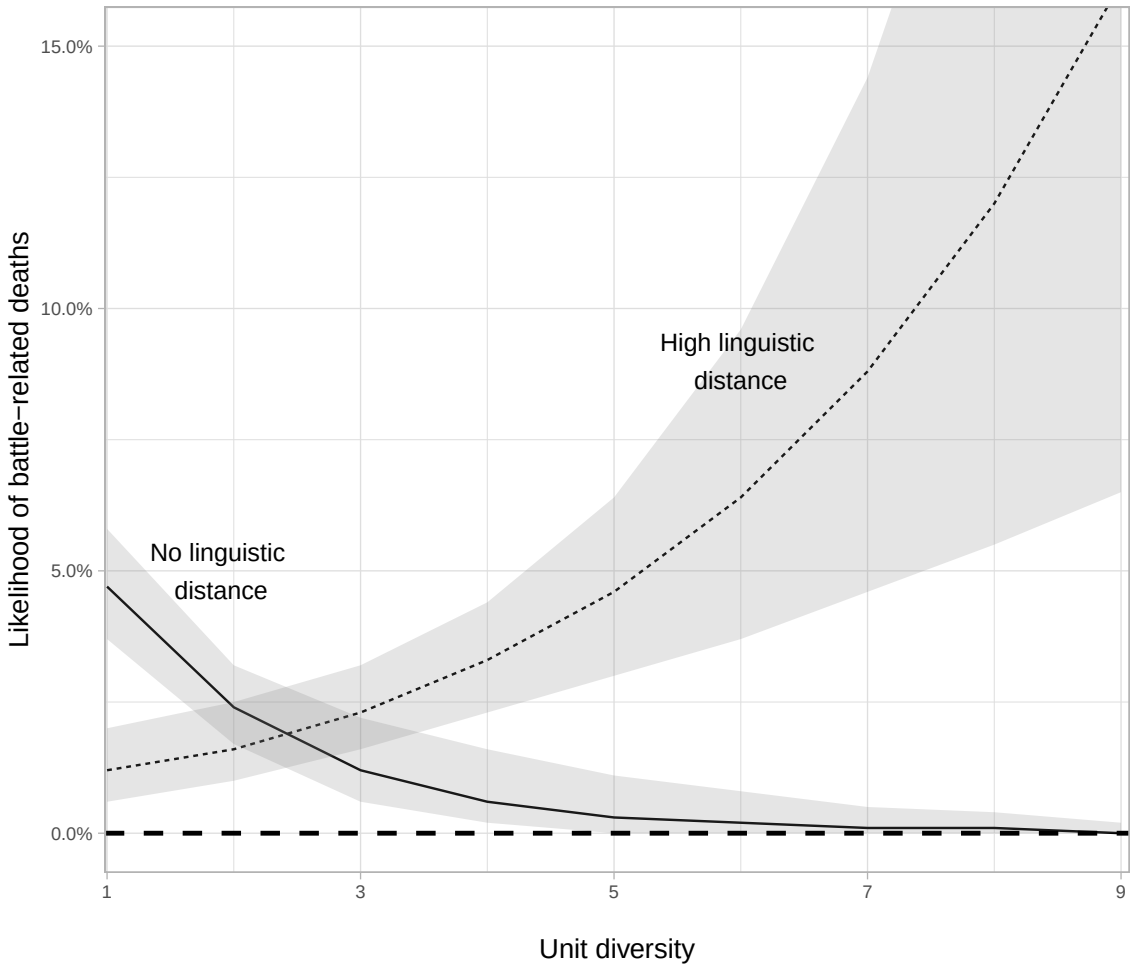
4.4.4 Example illustrations

In this section we provide anecdotes from 15 semi-structured interviews with former and active peacekeepers.[59] Table 11 on page 93 provides an overview of our sample. These in-depth qualitative accounts provide further evidence in favor of the mechanisms we lay out in our theory section. Specifically, they provide additional information about the conditions under which daily peacekeeping operations are af-

58. The figures show the effect of unit diversity even turning positive when there are high coordination barriers. Keeping in mind that these results hold "troop size" constant, this highlights an important policy trade-off: when a force commander is free to allocate new troops from another TCC, it seems to be more beneficial to boost an existing infantry contingent with more infantry contingents than to send additional unit types instead.

59. For exemplary questions that were designed to frame the conversation if necessary, please see the appendix.

Figure 15: Predicted probability of battle-related deaths: linguistic distances



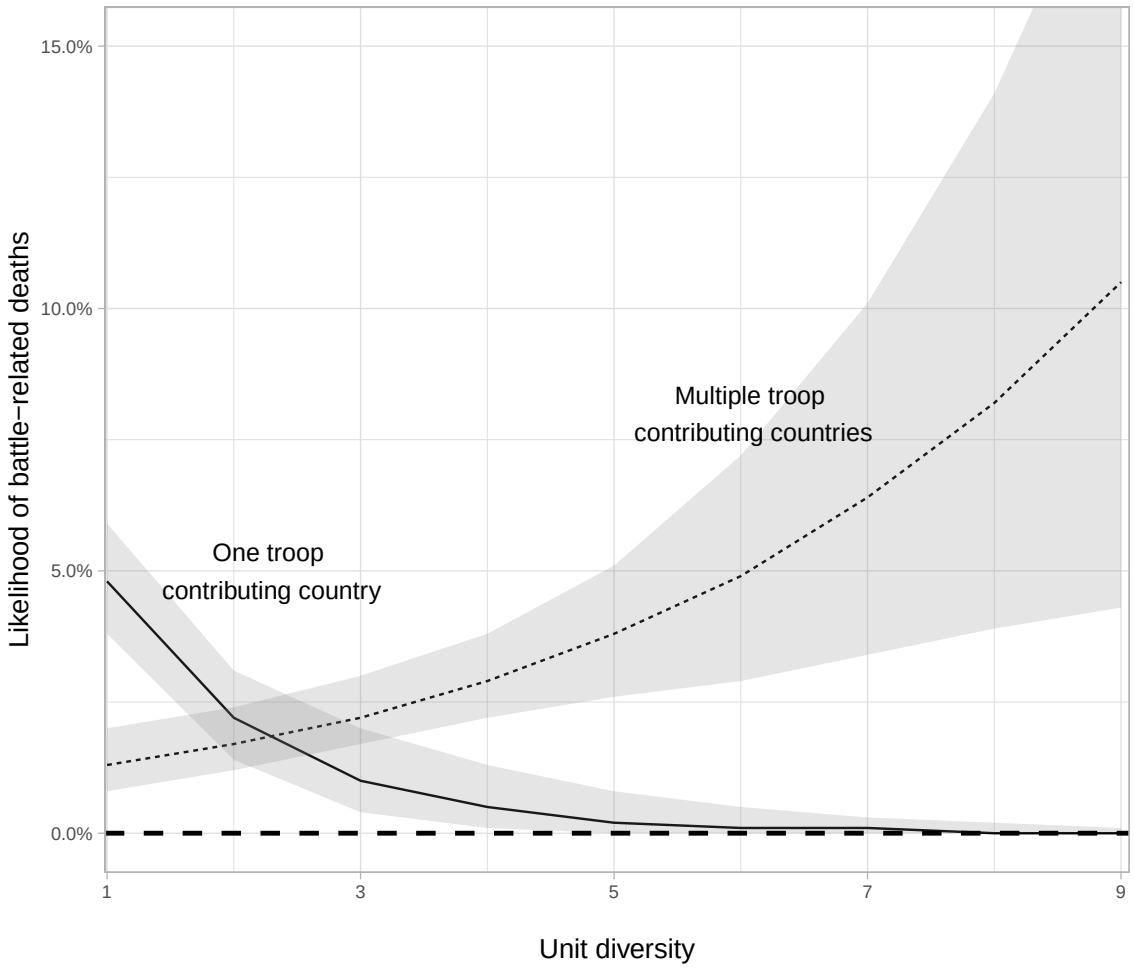
The figure is based on model 2. The 95% confidence intervals are based on non-parametric bootstraps with 2000 resamples. Predicted probabilities are calculated based on observed covariate values. Linguistic distance is set to 0 (no distance) and 1 (maximal distance) respectively.

fected by the two dimensions of troop diversity, *unit diversity* and *TCC diversity*. Evidence from these interviews suggest that unit diversity helps improve operations while TCC diversity can hamper them in the ways laid out in our theoretical discussion.

Unit diversity: specialization and division of labor

Evidence from interviews with peacekeepers provide support for our understanding that having a diverse set of units on the ground greatly increases peacekeepers' ability to cope with the security environment and to fulfill their mandate. One of our interviewees summarizes division of labor as follows: "Every contingent in the base has a specific task. Bangladesh has an aviation system that helps us with our tasks, like medical evacuation. The Argentinians [maintain] the hospital. The Indians were the infantry.

Figure 16: Predicted probability of battle-related deaths: TCC diversity



The figure is based on model 4. The 95% confidence intervals are based on non-parametric bootstraps with 2000 resamples. Predicted probabilities are calculated based on observed covariate values. "Same TCC" is set to 1 (same TCC) and 1 (multiple TCCs) respectively.

Table 11: Interviewee sample description

Ranks	Captain (5), Lt Colonel (4), Major (2), Lt General (1), Colonel (1), Lieutenant (1), Staff sergeant (1)
Nationalities	Brazil (6), Argentina (2), Nepal (2), Ghana (1), Germany (1), Norway (1), Netherlands (1), India (1)
Missions	MINUSTAH (4), MINUSCA (3), UNMISS (3), MINUSMA (2), UNMIS (1), UNIMSET (1), UNFICYP (1), UNIFIL (1), MONUC (1)
Appointments	Engineering (5), CIMIC/liason officer (4), military observer (4), operations/staff officer (4), UN instructor (3), platoon leader (2), translator (2), terrain analyst/intelligence officer (1), dep. chief of mission's CIMIC branch (1), mission section commander (1), military chief of staff (1), mission force commander (1)

We also had a Brazilian navy. The other Brazilians were the infantry, except our engineering company."

As commander, being able to draw on this unit diversity is paramount. At the highest level of mission

command, a former Military Chief of Staff who also has ISAF experience tells us: "Military specialized units, called enablers or tactical support units, are not only critical assets for maneuver forces, but have proven to play an increasingly important role in peace support, stability or COIN [counterinsurgency] operations."

Talking about the usefulness of specific unit types, interviewees explain how each unit type contributes to the fulfillment of the mandate. For example, one interviewee tells us that peacekeeping bases have a "golden circle" surrounding them, which is the area in which patrols are still allowed. The width of this circle is a function of the time that it takes to get back to the base in case of an emergency. He/she explains how having an aviation unit greatly increases this circle, boosting patrolling radii. Further, some units can be crucial for the continuation of daily operations and patrols. For example, one interviewee explains the importance of having engineering units on the base: "Engineers [are] very important in Africa, considering there are no roads and no machinery from the same country to make roads better or to clean them. During the rainy season it floods everything, so basically the roads almost disappear. But there is no machinery [...] to clean the roads." Another interviewee summarizes it as follows: "Though you might think that infantry, engineers, and aviation are different units, they in fact work together. [They are] pretty close to each other. It's only the infantry that physically goes out and patrols and tries to contain violence, but even when doing that there are other elements."

Moreover, several interviewees mention how lack of specialized units can tie up valuable and limited manpower. For example, one interviewee explains that a lot of troops were actually bound to provide convoy security. He/she continues "which is not really helpful [...] if your infantry capacity is already limited. [At some point] Senegal provided a convoy security [transportation] battalion, [...] which actually opened up capacity within the infantry companies to actually do what they needed to be doing. [...] If you have dedicated capacity to provide convoy security, your regular infantry troops can be used to do regular infantry tasks."

TCC diversity: friction and cooperation problems

Perceived or actual attitudinal differences in their approach to peacekeeping tasks, as well as linguistic and cultural differences among troops, have been raised by almost all of our interviewees as one of the obstacles to cooperation and coordination among units. Many of them recalled how daily operations could benefit from more effective cooperation among troops. The aforementioned Military Chief of Staff

put it this way: "One of the challenges within the mission [...] is the reluctance by default to share information with [...] other parts of the organization. For several, not good, reasons. And much of it is, to my experience, based on cultures and ways of conducting military affairs, which are not the way we would see as very efficient in a modern alliance like in NATO. The withholding of information may in some cases give you an upper hand in a situation. [...] Just finding ways of exchanging information [...] was actually one of the chief of staff's key functions. [...] Making sure that people had a certain level of collaboration between the different branches."

Further, one of our interviewees recalls how he/she had to travel with their Force Commander as a translator and further elaborates that the different accents of TCCs sometimes made communication very more difficult. The translator recalled one occasion when the Force Commander experienced unexpected communication problems and they were not traveling together, so that the Force Commander had to wait for his/her arrival on the scene in order to continue their work. This account received further corroboration in an interview with a Force Commander: "One [issue] is communication [equipment], the second is the language. Because although English is the [standard] language, but everybody is not fluent. The Chinese don't understand so they brought a liaison officer. So by the time you pass a message, it has to be understood, it has to be interpreted correctly, it has to be passed on to the troops on the ground, so that is a serious challenge. There is obviously a time delay."

Prompted on "communication between troops", another interviewee formerly deployed to Mali told us that "[...] command and control structures and how you communicate with each other on the battlefield, that is really problematic. First of all, a lot of the TCCs don't have effective equipment when they communicate with each other, then there is no compatibility within those communications systems, and then there are language issues in many cases. So that is really difficult, and what you see often is that on battalion-level, those are national assets. So a TCC [...] communicates [...] internally within their own means and capabilities. Similarly, one interviewee recalled that "Every TCC has different radios. [...] Communication is a big issue. Especially if you are patrolling in a dangerous area and you have to call for a Quick Reaction Force, or something like that [...]."

One of our interviewees explained how these differences significantly impacted daily operations. "Sometimes we had some problems with some countries because we felt that they were not willing so much to be operating and to be there. Sometimes we did not like to work with them. [...] We were not

willing to work with them because they did not like working. [...] It is a matter of culture. [...] Like the [country] troops, it was hard work with them. [...] They also had an engineering company in the field. When there was a request for the engineers to operate, they always had an excuse to not be in the field. We saw that they did not really want to engage in what they should be engaging. [...] When they received an order, they always had a lame excuse."

It is likely that mission leadership is also aware of existing or potential coordination problems among culturally diverse units. It is, however, not uncommon for very diverse troops to be deployed together since the UN has to work with limited resources and manpower to undertake enormous tasks in ongoing conflict situations. As the aforementioned Military Chief of Staff explained: "One of the elements of guidance I gave to my staff when they started looking at the options for a reconfiguration of the UNMISS force, was to look at how we can geographically reshuffle the mission so that we could have several units with at least certain closer cooperation. [...] Even these very simple ways of bringing the whole mission up to at least a perceived level of professionalism was quite a challenge." Ultimately, the plans to reshuffle the mission were not carried out. In another interview, a mission Force Commander confirmed that when it comes to the allocation of contingents, "we don't deploy based on nationality usually." When probed on the term "usually", he explained how sometimes the host country may have certain "sensitivities" based on the political environment, and veto the deployment of a specific TCC close to a specific country border. Based on these and similar accounts, we conclude that the data generating process of TCC diversity is not systematically related to previous violence levels.

4.5 Conclusion

This paper examines the effect of troop diversity on peacekeeping effectiveness. We argue that troop diversity occurs along two dimensions, unit diversity and TCC diversity, which interactively impact peacekeeping effectiveness. Specifically, we argue that unit diversity improves effectiveness in preventing battle-related violence through division of labor among peacekeeping troops. Unit diversity increases the range of equipment and skills available on the ground. With such comparative advantages, one contingent may specialize on logistics, while a second secures the camp, and a third better-equipped one patrols the wider vicinity. As a result, peacekeepers' ability to effectively fulfill their mandate increases.

TCC diversity, however, strongly moderates this effect, as cooperation and coordination are impeded linguistic and cultural barriers.

Using new subnational peacekeeping deployment data, Geo-PKO (Cil 2019), we find support for our theory. Each additional unit type in a given grid-month reduces the likelihood of battle-related deaths if peacekeepers come from the same TCC or experience no language difficulties. However, this negative effect of additional functionally specialized units on the likelihood of violence is lost when they come from different TCCs, and especially when linguistic distances are large.

Our study contributes to the academic literature on peacekeeping effectiveness, and may inform future policies on troop composition. First, the theoretical framework attempts to bridge the gap between peacekeeping and military studies. Increasingly, peacekeeping operations are tasked with robust and complex mandates in ongoing conflict environments. This renders certain aspects of peacekeeping, for better or worse, more comparable to counterinsurgencies. While our purpose is not to compare these types of operations, as comparisons between them have been made elsewhere (Howard 2019), we suggest that some of the existing military literature on force composition may be useful to better understand peacekeeping effectiveness. Second, our findings speak directly to existing peacekeeping scholarship. They build on previous work and extend it by acknowledging variation in peacekeeping troops. By disaggregating troops based on their specialization, i.e. unit type and origin, we offer a fine-grained understanding of local peacekeeping effectiveness. Finally, our results have policy implications for the design of future peacekeeping missions and their deployment patterns. Force commanders have to make several trade-offs when they are assigning units from different TCCs to a location. Naturally, they have to take into account their limited budget and availability of troops. Beyond these considerations however, our results show that adding specialized units in a deployment location is highly beneficial for the local performance of peacekeepers in general. When these functionally diverse troops are also culturally diverse however, coordination issues may impede effective cooperation and coordination between units.

There are several avenues for future research that follow from the theory and analysis presented here. Future research should examine the effect of force structure on other types of violence. For example, previous research highlights the importance of distance between peacekeepers and locals for the protection of civilians. Addressing one-sided violence requires other skills than preventing battle-

related violence (Cil 2019; Bove and Ruggeri 2016). More specialized units, moderated by troops' cultural distance to the local population, may impact the effectiveness in preventing or reducing one-sided violence. In addition, our proposed mechanisms are likely to have implications for peacekeepers' own security as well as their conduct. Newly available data on attacks against peacekeepers (Bromley 2018) can be used to test whether troop diversity help reduce attacks against peacekeepers, or whether certain troop compositions make peacekeepers more vulnerable to attacks. Similarly, troop diversity in a deployment location may deter misconduct by troops. Presence of a diverse set of TCCs may help in addressing sexual exploitation and abuse by providing more oversight and accountability. We also acknowledge that our identification strategy relies on conditioning on observables. Future studies may identify opportunities for causal identification that do not depend on the selection-on-observables assumption, and reexamine our findings of force structure and its effects on peacekeeping success. Finally, given the importance of force structure on peacekeeping effectiveness, future studies should examine the determinants of variation in subnational deployment in a more detailed manner.

4.6 Online Appendix

4.6.1 Appendix A: Examples of interview questions

The interviews were only loosely guided by these questions. More often, we had freely flowing conversations with our participants, and flexibly probed our mechanisms. We did this probing in a way as to minimize participant expectation, trying not to lead them. Participants were generally not informed about the specifics of our research question or theory *a priori*. Any information we obtained after disclosing our objectives, though informative, were not used in this paper.

1. Consent questions, in line with ethical compliance.
2. Background questions, recording information on participant's deployments.
3. "Tell us about a normal day during your deployment. What were your daily duties?"
4. "Were there any other [TCCs/units] deployed in your location? Did you interact with them? What were their duties?"
5. "Can you have a guess how it gets decided where specific [TCCs/units] are being sent within the mission?"
6. "Can you have a guess why there are sometimes more [TCCs/units] deployed to some locations than to others?"
7. "During daily duties or joint operations, how do you communicate between [TCCs/units]? Is that flat or through triangulation? [Explain if necessary]"
8. "Tell us about the hierarchy when there are multiple [TCCs/units] in one place. Who leads?"
9. "What does [specific unit type] do? How is it relevant to your mandate?"
10. "Do different TCCs play sports or games off-duty? Did you experience any language issues?"
11. "Are there some TCCs that were different from the others, maybe in their training or equipment?"

4.6.2 Appendix B: Supplementary results

Table 12 on the next page contains three additional sets of model, each including both of our moderating variables (TCCs' linguistic distance and our binary measure of TCC diversity). The first set of models does not include country fixed effects to assess model dependence and address the potential of overfitting. Due to the limited number of realizations of our binary dependent variable, the second set of models shows Firth rare events logistic regressions for bias adjustment (Firth 1993). Finally, we address the fact that there are no observations with eight or nine unique unit types when TCC diversity is low, which would normally lead to a funcitonal extrapolation for the construction of counterfactuals. Due to our employment of CBGPS weights, this should only pose a minor concern as functional form assumptions are relaxed. However as an additional check, the third set of models analyze a subsample which excludes observations with eight or nine unique unit types.[60] Across all of these models, our findings stay substantially and statistically significant.

Table 13 on page 102 shows results for religious distances as alternative moderating variable. Just as linguistic distances, religious distances are captured using a tree-based method with the distance between two countries being larger the more their main religions "branch out" from each other historically. This measure is also taken from (Spolaore and Wacziarg 2016). Not only does it constitute an alternative proxy for language difficulties, therefore serving as a reliability check, but it may also capture the cultural friction we encountered among peacekeepers even better than linguistic distances. The results corroborate our main findings: unit diversity decreases the likelihood of battle-related events if troops are similar in their contributing countries' main religions, however this effect is inverted when their TCCs are religiously dissimilar.

For further comparison and to motivate future research into this topic, Table 14 on page 103 lists models using violence against civilians ("one-sided violence") by rebel armed actors as outcome variable, while keeping the same independent variables as in our main models. We restrict this analysis to one-sided violence by rebels, because recent studies suggest that protection through peacekeepers tends to be limited when violence is being perpetrated by government forces (Fjelde, Hultman and Nilsson 2019; Carnegie and Mikulaschek 2017). In line with the findings of Cil (2019), which highlight the relevance of distinguishing between battle-related and one-sided violence when assessing different

60. We conduct a similar subsample analysis restricting the number of troops to only contain areas with less than 1000 personnel. Our findings are equally supported in this subsample.

forms of peacekeeping effectiveness, we do not find evidence that military specialization and coordinated force projection improve the prevention of one-sided violence.

Table 12: Alternative specifications

	Occurrence of battle-related deaths					
	No FE	*No* FE	*Firth* logit	*Firth* logit	*Sub-*sample	*Sub-*sample
Unit diversity (t-1)	−0.854***	0.330	−0.836***	0.323	−0.910***	0.325
	(0.240)	(0.226)	(0.223)	(0.214)	(0.287)	(0.226)
TCCs' linguistic distances (t-1)	−3.101***		−3.054***		−3.134***	
	(1.163)		(1.106)		(1.204)	
Same TCC (t-1)		2.961***		2.888***		2.888***
		(0.992)		(0.944)		(1.022)
Battle deaths(bin)(3m ave.)	1.267***	1.243***	1.244***	1.220***	1.153***	1.125***
	(0.317)	(0.331)	(0.304)	(0.318)	(0.332)	(0.347)
Neighb. battle deaths(bin)(t-2)	1.519***	1.543***	1.500***	1.523***	1.603***	1.629***
	(0.259)	(0.258)	(0.248)	(0.246)	(0.294)	(0.300)
Number of troops(ln)(t-2)	0.488***	0.469***	0.474***	0.456***	0.462***	0.441***
	(0.126)	(0.125)	(0.119)	(0.118)	(0.129)	(0.128)
Neighb. Troops(ln)(t-2)	0.052	0.052	0.051	0.051	0.058	0.053
	(0.065)	(0.065)	(0.061)	(0.061)	(0.069)	(0.070)
Mountainous Terrain	0.356	0.341	0.336	0.318	0.813	0.781
	(0.596)	(0.611)	(0.564)	(0.576)	(0.858)	(0.863)
Nightlights Coverage	7.040*	6.591	7.112**	6.675*	4.515	4.974
	(3.804)	(4.168)	(3.561)	(3.905)	(7.729)	(8.324)
Excluded Ethnic Groups	−0.122	−0.162	−0.125	−0.163	−0.287	−0.319
	(0.204)	(0.208)	(0.189)	(0.193)	(0.252)	(0.255)
Distance to Capital	−0.249*	−0.268**	−0.247*	−0.265**	−0.398	−0.394
	(0.137)	(0.137)	(0.131)	(0.130)	(0.378)	(0.389)
Distance to Border	−0.162	−0.144	−0.160	−0.144	−0.218	−0.203
	(0.121)	(0.117)	(0.113)	(0.110)	(0.160)	(0.160)
Unit diversity * Ling. dist.	1.295***		1.272***		1.377***	
	(0.398)		(0.374)		(0.446)	
Unit diversity * Same TCC		−1.318**		−1.276**		−1.340***
		(0.407)		(0.380)		(0.453)
Constant	−3.433***	−6.024***	−3.359***	−5.912***	−2.178	−4.837**
	(1.295)	(1.137)	(1.235)	(1.074)	(2.705)	(2.420)
Observations	4660	4660	4660	4660	4613	4613
CBGPS matching	✓	✓	✓	✓	✓	✓
Country FE	X	X	X	X	✓	✓
Time cubic polynomial	✓	✓	✓	✓	✓	✓
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 13: Religious distances

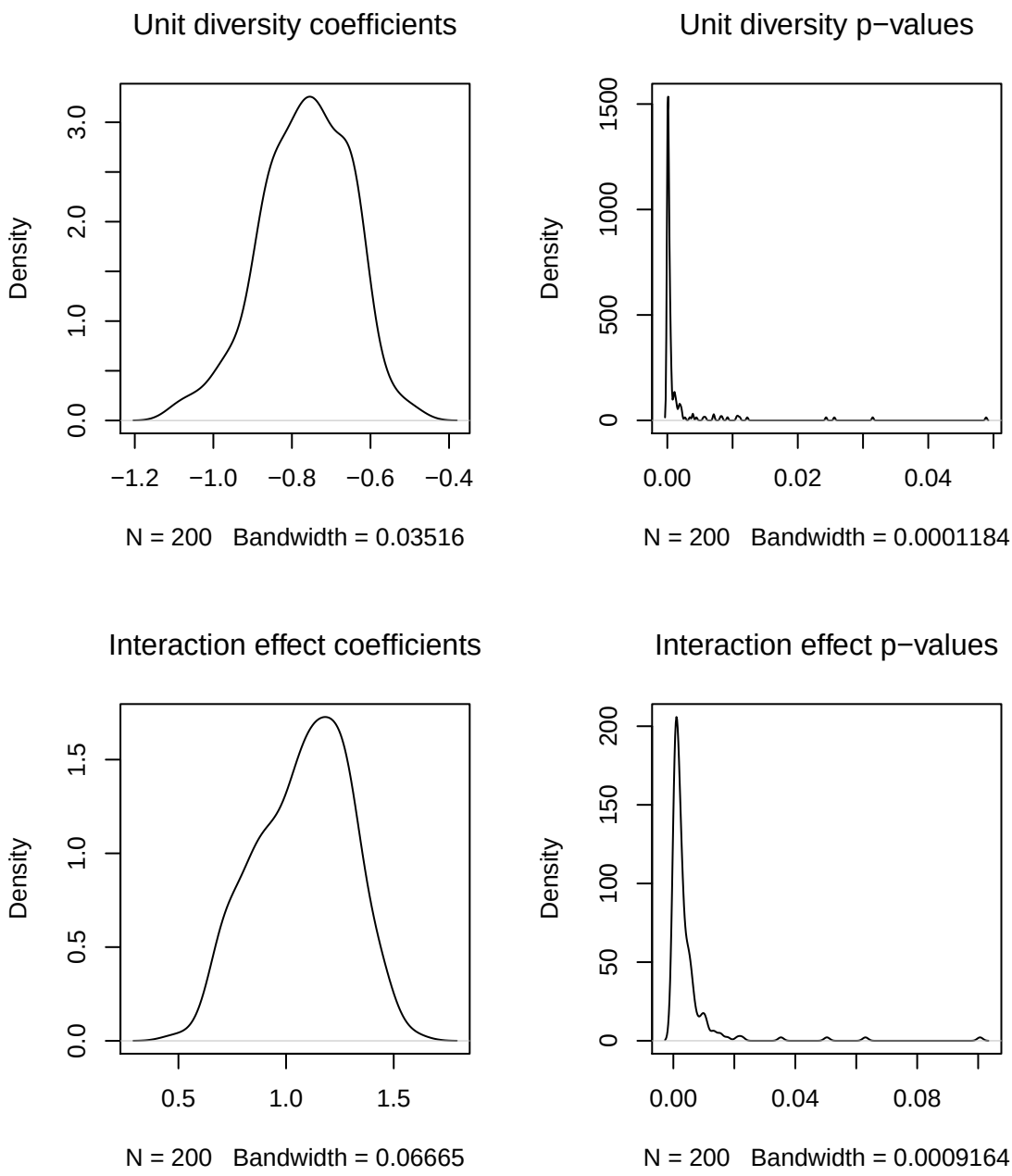
	Occurrence of battle-related deaths	
	(1)	(2)
Unit diversity (t-1)	-0.895^{***}	-0.917^{***}
	(0.230)	(0.260)
TCCs' religious distances (t-1)	-3.643^{***}	-3.506^{***}
	(1.201)	(1.192)
Battle deaths(bin)(3m ave.)	1.264^{***}	1.155^{***}
	(0.312)	(0.329)
Neighb. battle deaths(bin)(t-2)	1.500^{***}	1.602^{***}
	(0.254)	(0.293)
Number of troops(ln)(t-2)	0.507^{***}	0.496^{***}
	(0.131)	(0.134)
Neighb. Troops(ln)(t-2)	0.057	0.060
	(0.063)	(0.068)
Mountainous Terrain	0.264	0.668
	(0.588)	(0.885)
Nightlights Coverage	7.294^{*}	5.736
	(3.762)	(7.953)
Excluded Ethnic Groups	-0.138	-0.252
	(0.210)	(0.251)
Distance to Capital	-0.265^{*}	-0.373
	(0.139)	(0.381)
Distance to Border	-0.172	-0.247
	(0.119)	(0.158)
Unit diversity * Relig. dist.	1.518^{***}	1.511^{***}
	(0.425)	(0.439)
Constant	-3.399^{***}	-2.419
	(1.287)	(2.734)
Observations	4660	4660
CBGPS matching	✓	✓
Country FE	X	✓
Time cubic polynomial	✓	✓
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 14: Effect on one-sided violence by rebels

	Occurrence of one-sided violence by rebels	
	(1)	(2)
Unit diversity (t-1)	0.429**	0.224
	(0.195)	(0.142)
TCCs' linguistic distances (t-1)	0.507	
	(0.574)	
Same TCC (t-1)		−0.368
		(0.503)
Battle deaths(bin)(3m ave.)	1.332***	1.333***
	(0.422)	(0.422)
Neighb. battle deaths(bin)(t-2)	2.115***	2.113***
	(0.367)	(0.367)
Number of troops(ln)(t-2)	0.110	0.113
	(0.133)	(0.133)
Neighb. Troops(ln)(t-2)	−0.011	−0.011
	(0.061)	(0.060)
Mountainous Terrain	1.617**	1.626**
	(0.670)	(0.671)
Nightlights Coverage	−11.908	−11.985
	(9.879)	(9.884)
Excluded Ethnic Groups	0.350	0.348
	(0.290)	(0.290)
Distance to Capital	−0.550*	−0.552*
	(0.307)	(0.304)
Distance to Border	−0.515***	−0.519***
	(0.127)	(0.128)
Unit diversity * Ling. dist.	−0.242	
	(0.230)	
Unit diversity * Same TCC		0.181
		(0.208)
Constant	−1.041	−0.617
	(2.317)	(2.465)
Observations	4660	4660
CBGPS weights	✓	✓
Country FE	✓	✓
Time cubic polynomial	✓	✓
Note:		*p<0.1; **p<0.05; ***p<0.01

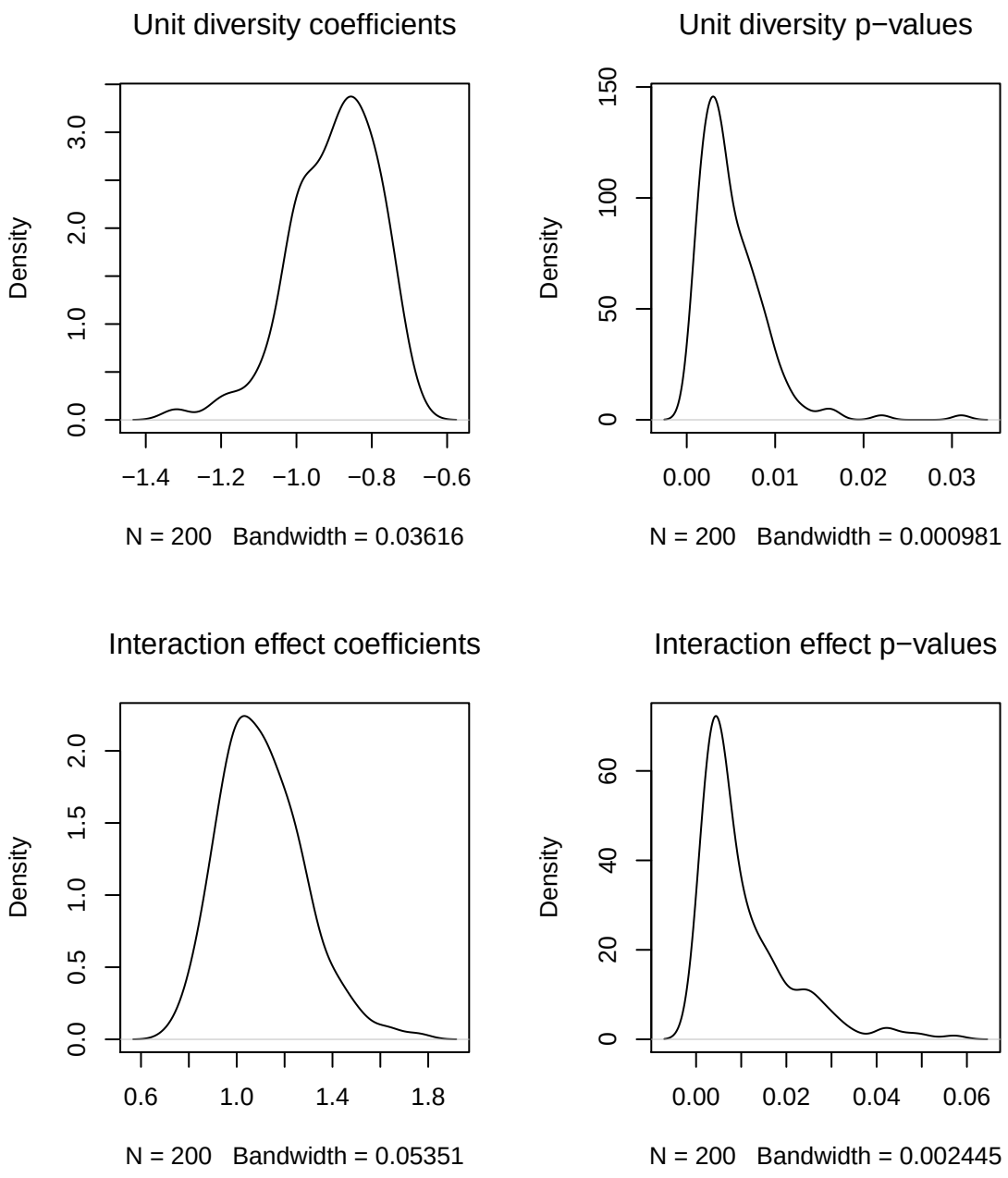
Figure 17 on the following page and Figure 18 on page 106 show simulation results. These serve as a bounding exercise to learn about the sensitivity of our results. As explained above, some areas may have a higher or lower number of functional units due to previous conflict activity. While we address this issue of reversed causality through several channels, these simulations help to better understand the degree to which such a reversed relationship may drive our results. For each simulated dataset, we raise or attenuate the treatment variable (assign one functional unit more or less) of a a random sample of 40% of those observations that have experienced violence in the previous three months. In other words, we assume that a share of the data we observe was shaped by reversed causality and manipulate the data to alter that share. We did the same with 20% of the observations, however here we only show the figures of the "more extreme" 40% for space reasons. Concentrating on the two most important effect sizes, the coefficients of unit diversity and of the interaction effect, the simulated results stay robust. While this does not alleviate concerns over endogeneity, it provides further corroboration that we have identified a systematic pattern.

Figure 17: Simulation results: endogenously increasing unit diversity



The figure is based on model 2 in the paper: unit diversity (t-1) is interacted with linguistic distances, employing CBGPS matching, country fixed effects, and time cubic polynomials. The top row shows the coefficients and p-values of 200 unit diversity coefficients, while the bottom row shows the coefficients and p-values of 200 interaction effects. Each of the 200 analyses is based on newly simulated data. In each iteration, the data adjust a random sample of 40% of the potentially endogeneous observations (violence was observed in previous three months) by raising the number of unit types by one additional unit.

Figure 18: Simulation results: endogenously attenuating unit diversity



The figure is based on model 2 in the paper: unit diversity (t-1) is interacted with linguistic distances, employing CBGPS matching, country fixed effects, and time cubic polynomials. The top row shows the coefficients and p-values of 200 unit diversity coefficients, while the bottom row shows the coefficients and p-values of 200 interaction effects. Each of the 200 analyses is based on newly simulated data. In each iteration, the data adjust a random sample of 40% of the potentially endogeneous observations (violence was observed in previous three months) by lowering the number of unit types by one additional unit.

5 Conclusion

How does the composition of security forces in counterinsurgency and peacekeeping operations influence their behaviour and performance? In this thesis I find that the composition of security forces has salient implications for their effectiveness and loyalty. In my first paper I show that a more differentiated security apparatus is more likely to win counterinsurgency wars: the fragmentation of the security forces allows them to specialise on distinct tasks and engage in a division of labour, rendering them more effective. This effect however is strongly moderated by coup risk due to the ambivalent nature of fragmentation. Under high coup risk, a fragmented military is rather indicative of coup-proofing efforts aimed at rendering the military less effective.

In my second paper I draw on this distinction in conceptualising security force fragmentation, showing how differentiation leads to a lower probability of security force defections due to military personnel bandwagoning with the government side. As expected however, fragmentation under high coup risk leads them to bandwagon with the opposition. In the third paper my co-author and I find that peacekeepers also benefit from specialisation and a division of labour. Maintaining different unit types in an area renders the peacekeepers more effective in deterring violence between combatants. However this beneficial effect is strongly moderated by the linguistic and cultural friction between the troops, impeding coordination of their actions.

5.1 Research contribution and impact

My findings have multiple academic and policy implications. The first two articles shed light on the relevance of considering security forces' composition and command structure, as well as the role of appropriate training and the double-edged effects of coup-proofing. These insights support policy-makers who seek to build counterinsurgency capacity. They also help external observers to better understand the trade-offs that especially autocrats face when dealing with a popular uprising, allowing for a more nuanced assessment when deciding on whether and when to intervene. Academically, the findings augment the core of civil-military studies. They provide an important step towards resolving the discrepancy and contradicting assumptions in the existing literature on the relationship between counterbalancing, differentiation, and the coup-civil war trap (e.g. J. M. Powell 2012b versus Pilster, Böhmelt and Tago

2016). They also constitute the first systematic evaluation of the role of security forces' composition in their decision to bandwagon with the winning side during an uprising.

The third article suggests crucial implications for the design of future peacekeeping missions and their deployment patterns. There are several trade-offs involved when assigning units from different contributing countries to a location, including a limited budget and the availability of troops. Beyond these considerations however, my results indicate that additional specialised units in a deployment location are highly beneficial for the local performance of peacekeepers. When these functionally diverse troops are also culturally diverse however, coordination issues may impede effective cooperation and coordination between units. As an important academic contribution, the theoretical framework bridges the gap between peacekeeping and military studies. Peacekeeping operations are increasingly tasked with robust mandates in active conflict zones. This renders peacekeeping, for better or worse, more comparable to counterinsurgencies. My theoretical framework suggests that scholarship on the force composition in peacekeeping operations can draw important lessons from findings in the military studies literature. In addition, the theory and findings speak directly to existing peacekeeping scholarship. They extend it by acknowledging previously unexplored variation in peacekeeping troops: by disaggregating based on unit types, the article offers a fine-grained understanding of local peacekeeping effectiveness.

Combining different strands of literature, and tying them together with my novel findings on the implications of security force composition, this thesis contributes to core scholarship on military studies, civil-military relations, and peacekeeping. My results yield relevant policy implications for force commanders of state militaries and UN missions, providing suggestions on force and command structures. They also indicate important dynamics based on which the international security community may adjust their risk assessments and threat evaluations.

5.2 Avenues for future research

There are several avenues for future research to build on and extend this thesis. First, measuring "counterinsurgency effectiveness" is a contested endeavour, and future research may replicate my results using other outcome variables. Second, my conceptualisation of differentiation suggests that counterbalancing, other than differentiation, is associated with the security forces being kept close to the nation's capital and that they are deployed less to remote locations (J. Powell 2019; J. M. Powell 2012b). This

would be associated with a strong negative correlation between military effectiveness and proximity to the capital. While this correlation is always probable (see also Schutte 2017), based on my framework I expect it to be significantly stronger under high coup risk than under low coup risk. Third, the mechanisms I identify in my first two papers may also be applicable to military effectiveness in inter-state conflict, and such an extension of my research would seem worthwhile. Fourth, while the third paper benefits from very detailed geo-spatial data, this is not the case for the first and second paper. The extent to which horizontal fragmentation of the security forces can serve as a proxy for specialisation is questionable, and I expect estimation efficiency to suffer. This is especially relevant given the disputed data quality of the IISS (De Bruin 2020b). Therefore, testing more empirical implications of my theoretical models and conducting more qualitative work are worthwhile.

When it comes to peacekeeping, future research may focus on the effect of other aspects of force structure on other types of violence: violence against civilians is an important aspect in peacekeeping, and recent research suggests that peacekeepers' cultural distance to locals is an important factor preventing it. Interacting this with functional diversity may yield important insights into peacekeeping dynamics. In addition, the proposed mechanisms are likely to have implications for peacekeepers' own security. Troop diversity may help to reduce attacks against peacekeepers, and newly available data on attacks against peacekeepers would make such an investigation possible (Bromley 2018). Moreover, the third article is restricted to UN peacekeeping, and from our analysis it is unclear whether the same pattern applies to peacekeeping by, for example, the African Union.

All three studies employ identification strategies that rely on conditioning on observables. Future research should aim to use research designs, or methods of causal inference, that account for unobserved confounding. Finally, building on my findings, more predictive exercises may contribute to inform policy makers on expected patterns of effectiveness.

References

Acemoglu, Daron, and James A. Robinson. 2001. “A Theory of Political Transitions”. *American Economic Review* 91 (4): 938–963.

Agresti, Alan. 2002. *Categorical Data Analysis.* New Jersey: Wiley.

Ahmed, Ishtiaq. 1996. *State, Nation and Ethnicity in Contemporary South Asia.* London: Pinter.

Akcinaroglu, Seden, and Elizabeth Radziszewski. 2013. “Private Military Companies, Opportunities, and Termination of Civil Wars in Africa”. *Journal of Conflict Resolution* 57 (5): 795–821.

Ari, Erkan, and Yildiz Zeki. 2014. “Parallel Lines Assumption in Ordinal Logistic Regression and Analysis Approaches”. *International Interdisciplinary Journal of Scientific Research* 1 (3): 8–23.

Auerswald, David P., and Stephen M. Saideman. 2014. *NATO in Afghanistan: Fighting together, fighting alone.* Princeton: Princeton University Press.

Autesserre, Séverine. 2014. *Peaceland: Conflict resolution and the everyday politics of international intervention.* Cambridge: Cambridge University Press.

Avant, Deborah D. 2007. “Political Institutions and Military Effectiveness: Contemporary United States and United Kingdom”. In *Creating Military Power: The Sources of Military Effectiveness,* edited by Risa Brooks and Elizabeth Stanley. Stanford: Stanford University Press.

Axelrod, Robert M. 1984. *The Evolution of Cooperation.* Renewed 20. New York: Basic Books.

Barany, Zoltan. 2011. “Comparing the Arab Revolts: The Role of the Military”. *Journal of Democracy* 22 (4): 24–35.

Belkin, Aaron, and Evan Schofer. 2003. “Toward a Structural Understanding of Coup Risk”. *Journal of Conflict Resolution* 47 (5): 594–620.

———. 2005. “Coup Risk, Counterbalancing, and International Conflict”. *Security Studies* 14 (1): 140–177.

Bell, Curtis, and Jun Koga Sudduth. 2017. “The Causes and Outcomes of Coup during Civil War”. *Journal of Conflict Resolution* 61 (7): 1432–1455.

Berman, Eli, Joseph H. Felter and Jacob N. Shapiro. 2018. *Small Wars, Big Data: The Information Revolution in Modern Conflict.* Princeton: Princeton University Press.

Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle.* Princeton: Princeton University Press.

Biddle, Stephen, and Stephen Long. 2004. “Democracy and Military Effectiveness: A Deeper Look”. *Journal of Conflict Resolution* 48 (4): 525–546.

Böhmelt, Tobias, and Govinda Clayton. 2018. “Auxiliary Force Structure: Paramilitary Forces and Progovernment Militias”. *Comparative Political Studies* (Comparative Political Studies) 51 (2): 197–237.

Böhmelt, Tobias, Christoph Dworschak, Ulrich Pilster and Julian Walterskirchen. 2020. “A Cross-National Analysis of Forced Population Resettlement in Counterinsurgency Campaigns”. *Peace Economics, Peace Science and Public Policy* 26 (1).

Böhmelt, Tobias, Andrea Ruggeri and Ulrich Pilster. 2017. “Counterbalancing, Spatial Dependence, and Peer Group Effects”. *Political Science Research and Methods* 5 (2): 221–239.

Bove, Vincenzo, Chiara Ruffa and Ana Ruggeri. 2020. *Composing Peace: Mission Composition in UN Peacekeeping.* Oxford: Oxford University Press.

Bove, Vincenzo, and Andrea Ruggeri. 2016. “Kinds of Blue: Diversity in UN Peacekeeping Missions and Civilian Protection”. *British Journal of Political Science* 46 (03): 681–700.

———. 2019. “Peacekeeping Effectiveness and Blue Helmets’ Distance from Locals”. *Journal of Conflict Resolution* 63 (7): 1630–1655.

Boyle, Michael. 2008. “Terrorism and Insurgency”. In *Contemporary Security and Strategy,* edited by Craig A Snyder. New York: Palgrave Macmillan.

Branch, Daniel. 2009. *Defeating Mau Mau, creating Kenya: Counterinsurgency, Civil War, and Decolonization.* New York: Cambridge University Press.

Brar, K S. 2000. *Operation Blue Star: The True Story.* New Delhi: UBS Publishers. *7th edition.*

Brinton, Clarence Crane. 1938. *The Anatomy of Revolution.* New York: Vintage Books. *Revised edition: 1965.*

Bromley, Sara Lindberg. 2018. “Introducing the UCDP Peacemakers at Risk dataset, sub-Saharan Africa, 1989–2009”. *Journal of Peace Research* 55 (1): 122–131.

Brooks, Risa A. 2003. “Making Military Might: Why Do States Fail and Succeed? A Review Essay”. *International Security* 28 (2): 149–191.

———. 2006. “An Autocracy at War: Explaining Egypt’s Military Effectiveness, 1967 and 1973”. *Security Studies* 15 (3): 396–430.

———. 2008. *Shaping strategy: The civil-military politics of strategic assessment.* Princeton: Princeton University Press.

———. 2013. “Abandoned at the Palace: Why the Tunisian Military Defected from the Ben Ali Regime in January 2011”. *Journal of Strategic Studies* 36 (2): 205–220.

———. 2017. *Military Defection and the Arab Spring.* Oxford Research Encyclopedia of Politics: Oxford University Press.

Brovkin, Vladimir N. 1994. *Behind the Front Lines of the Civil War: Political Parties and Social Movements in Russia, 1918-1922.* Princeton: Princeton University Press.

Burke, Jason. 2017. *Mutinous Ivory Coast soldiers seize control of three cities.* The Guardian. `https://www.theguardian.com/world/2017/jan/06/mutinous-ivory-coast-soldiers-seize-control-of-three-cities`. *Retrieved on: 23.11.2017.*

Callwell, Charles E. 1996. *Small Wars: Their Principles and Practice.* Lincoln: University of Nebraska Press.

Carey, Sabine C, Michael P Colaresi and Neil J Mitchell. 2016. “Risk Mitigation, Regime Security, and Militias: Beyond Coup-proofing”. *International Studies Quarterly* 60 (1): 59–72.

Carnegie, Allison, and Christoph Mikulaschek. 2017. “The Promise of Peacekeeping: Protecting Civilians in Civil Wars”. *SSRN Electronic Journal.*

Carter, David B., and Curtis S. Signorino. 2010. “Back to the Future: Modeling Time Dependence in Binary Data”. *Political Analysis* 18 (3): 271–292.

Caverley, Jonathan D, and Todd S Sechser. 2017. “Military Technology and the Duration of Civil Conflict”. *International Studies Quarterly* 61 (3): 704–720.

Celeski, Joseph D. 2009. *Policing and Law Enforcement in COIN - The Thick Blue Line.* JSOU Report 09-2: Joint Special Operations University.

Chen, Daniel L. 2016. *The Deterrent Effect of the Death Penalty? Evidence from British Commutations During World War I.* Social Science Research Network: Electronic Journal. `https://www.ssrn.com/abstract=2816255`. *Retrieved on: 21.11.2017.*

Chenoweth, Erica, and Orion A. Lewis. 2013. “Unpacking nonviolent campaigns”. *Journal of Peace Research* 50 (3): 415–423.

Chenoweth, Erica, and Maria J. Stephan. 2013. *Why Civil Resistance Works : the Strategic Logic of Nonviolent Conflict.* New York: Columbia University Press.

Cil, Deniz. 2019. “Who is Keeping the Peace? Local Peacekeeping Unit Composition and Effect-iveness”. *International Studies Association Annual Convention, Toronto, CA. Working paper.*

Cil, Deniz, Hanne Fjelde, Lisa Hultman and Desirée Nilsson. 2020. “Mapping blue helmets: Introducing the Geocoded Peacekeeping Operations (Geo-PKO) dataset”. *Journal of Peace Research* 57 (2): 360–370.

Clodfelter, Micheal. 2008. *Warfare And Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1494- 2007.* Jefferson, North Carolina: McFarland & Company, Inc.

Collier, Paul, V L Elliott, Havard Hegre, Anke Hoeffler, Marta Reynal-Querol and Nicholas Sambanis. 2003. *Breaking the Conflict Trap: Civil War and Development Policy (World Bank Report).* Washington DC: The World Bank / Oxford University Press.

Collier, Paul, and Anke Hoeffler. 2004. “Greed and Grievance in Civil War”. *Oxford Economic Papers* 56 (1): 563–595.

Costalli, Stefano. 2014. “Does peacekeeping work? A disaggregated analysis of deployment and violence reduction in the bosnian war”. *British Journal of Political Science* 44 (2): 357–380.

Cunningham, David E., Kristian S. Gleditsch and Idean Salehyan. 2009. “It Takes Two: A Dyadic Analysis of Cvil War Duration and Outcome”. *Journal of Conflict Resolution* 53 (4): 570–597.

Dahl, Marianne. 2016a. “Pockets of Opportunity: State Dissident Interactions and Regime Change”. *PhD Thesis at the Norwegian University of Science and Technology* (PhD Thesis).

———. 2016b. “Policy Brief: Military Defection”. *Peace Research Institute Oslo,* no. 18.

De Bruin, Erica. 2018. “Preventing Coups d’état”. *Journal of Conflict Resolution* 62 (7): 1433–1458.

———. 2020a. *How to prevent coups d’etat: Counterbalancing and regime survival.* New York: Cornell University Press.

———. 2020b. “Mapping coercive institutions: The State Security Forces dataset, 1960–2010”. *Journal of Peace Research* Online fir:002234332091308.

Derpanopoulos, George, Erica Frantz, Barbara Geddes and Joseph Wright. 2016. “Are coups good for democracy?” *Research & Politics* 3 (1): 1–7.

Downes, Alexander B. 2007. “Draining the Sea by Filling the Graves: Investigating the Effectiveness of Indiscriminate Violence as a Counterinsurgency Strategy”. *Civil Wars* 9 (4): 420–444.

Doyle, Michael W., and Nicholas Sambanis. 2006. *Making war and building peace: United Nations peace operations.* Princeton and Oxford: Princeton University Press.

Dworschak, Christoph. 2019. “Divide and Conquer: Military Differentiation and Counterinsurgency Effectiveness”. *Social Science Research Network.*

———. 2020. “Jumping on the Bandwagon: Differentiation and Security Defection during Conflict”. *Journal of Conflict Resolution. Online first.*

Dwyer, Maggie. 2015. “Tactical Communication: Mutiny as a Dialogue in West and Central Africa”. *Africa Spectrum* 50 (1): 5–23.

———. 2017. *Soldiers in Revolt: Army Mutinies in Africa.* London: C. Hurst & Co Publishers Ltd.

Escribà-Folch, Abel, Tobias Böhmelt and Ulrich Pilster. 2019. “Authoritarian regimes and civil–military relations: Explaining counterbalancing in autocracies”. *Conflict Management and Peace Science:* 073889421983628.

Fearon, James D. 1995. “Rationalist Explanations for War”. *International Organization* 49 (3): 379–414.

Fiala, Nathan, and Stergio Skaperdas. 2011. “Economic Perspectives on Civil Wars”. In *The Handbook on the Political Economy of War,* edited by Christopher J Coyne and Rachel L Mathers. Cheltenham: Edward Elgar Publishing Ltd.

Firth, David. 1993. “Bias reduction of maximum likelihood estimates”. *Biometrika* 80 (1): 27–38.

Fjelde, Hanne, Lisa Hultman and Desirée Nilsson. 2019. “Protection Through Presence: UN Peacekeeping and the Costs of Targeting Civilians”. *International Organization* 73 (1): 103–131.

Fong, Christian, Chad Hazlett and Kosuke Imai. 2018. “Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements”. *The Annals of Applied Statistics* 12 (1): 156–177.

Fortna, Virginia Page. 2008. *Does peacekeeping work? Shaping belligerents’ choices after civil war.* Princeton: Princeton University Press.

Friis, Karsten. 2010. “Peacekeeping and Counter-insurgency – Two of a Kind?” *International Peacekeeping* 17 (1): 49–66.

Galula, David. 1964. *Counterinsurgency Warfare: Theory and Practice.* Westport, Conneticut: Praeger Security International.

Gana, Nouri. 2013. *The Making of the Tunisian Revolution: Contexts, Architects, Prospects.* Edinburgh: Edinburgh University Press.

Gates, Scott. 2002. “Recruitment and Allegiance: the Microfoundations of Rebellion”. *Journal of Conflict Resolution* 46 (1): 111–130.

Gates, Scott, and Kaushik Roy. 2016. *Unconventional Warfare in South Asia: Shadow Warriors and Counterinsurgency.* London: Routledge.

Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. “Autocratic Breakdown and Regime Transitions: A New Data Set”. *Perspectives on Politics* 12 (2): 313–331.

Gentile, Gian P. 2013. *Wrong Turn: America’s Deadly Embrace of Counterinsurgency.* New York: The New Press.

Gill, Jeff. 1999. “The Insignificance of Null Hypothesis Significance Testing”. *Political Research Quarterly* 52 (3): 647–674.

———. 2018. “Editorial”. *Political Analysis* 26 (1): 1–2.

Gilligan, Michael J., and Ernest J. Sergenti. 2008. “Do UN interventions cause peace? Using matching to improve causal inference”. *Quarterly Journal of Political Science* 3 (2): 89–122.

Gilligan, Michael, and Stephen John Stedman. 2003. “Where Do the Peacekeepers Go?” *International Studies Review* 5 (4): 37–54.

Gleditsch, Kristian Skrede, Idean Salehyan and Kenneth Schultz. 2008. “Fighting at Home, Fighting Abroad: How Civil Wars Lead to International Disputes”. *Journal of Conflict Resolution* 52 (4): 479–506.

Goswami, Namrata. 2015. *Indian National Security and Counter-Insurgency: The use of force vs non-violent response.* New York: Routledge.

Haass, Felix, and Nadine Ansorg. 2018. “Better peacekeepers, better protection? Troop quality of United Nations peace operations and violence against civilians”. *Journal of Peace Research* 55 (6): 742–758.

Heng, Yee-Kuang. 2003. *War as Risk Management.* London School of Economics and Political Science: PhD Thesis.

Ho, Daniel, Kosuke Imai, Gary King and Elizabeth Stuart. 2007. “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference”. *Political Analysis* 15 (1): 199–236.

Howard, Lise M. 2019. *Power in Peacekeeping.* Cambridge: Cambridge University Press.

Howard, Lise Morjé, and Anjali Kaushlesh Dayal. 2018. “The Use of Force in UN Peacekeeping”. *International Organization* 72 (1): 71–103.

Howard, Lise Morjé, and Alexandra Stark. 2017. “How Civil Wars End: The International System, Norms, and the Role of External Actors”. *International Security* 42 (3): 127–171.

Hultman, Lisa, Jacob D. Kathman and Megan Shannon. 2016. “United Nations peacekeeping dynamics and the duration of post-civil conflict peace”. *Conflict Management and Peace Science* 33 (3): 231–249.

Hultman, Lisa, Jacob Kathman and Megan Shannon. 2013. “United Nations Peacekeeping and Civilian Protection in Civil War”. *American Journal of Political Science* 57 (4): 875–891.

———. 2014. “Beyond keeping peace: United nations effectiveness in the midst of fighting”. *American Political Science Review* 108 (4): 737–753.

Huntington, Samuel P. 1957. *The Soldier and the State: The Theory and Politics of Civil-Military Relations.* Cambridge, MA: Belknap Press.

Imai, Kosuke, and Marc Ratkovic. 2014. “Covariate balancing propensity score”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76 (1): 243–263.

International Institute for Strategic Studies. 2016. *The Military Balance.* London: International Institute for Strategic Studies.

Joes, Anthony James. 2011. “What is Guerilla Warfare?” In *The Handbook on the Political Economy of War,* edited by Christopher J Coyne and Rachel L Mathers. Cheltenham: Edward Elgar Publishing Ltd.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War.* Cambridge: Cambridge University Press.

Kapur, S. Paul. 2008. “Ten Years of Instability in a Nuclear South Asia”. *International Security* 33 (2): 71–94.

Kathman, Jacob D. 2013. “United Nations peacekeeping personnel commitments, 1990–2011”. *Conflict Management and Peace Science* 30 (5): 532–549.

Kathman, Jacob D., and Reed M. Wood. 2014. “Stopping the Killing During the “Peace”: Peacekeeping and the Severity of Postconflict Civilian Victimization”. *Foreign Policy Analysis* 12 (2): n/a–n/a.

King, Gary, and Langche Zeng. 2001. “Explaining Rare Events in International Relations”. *International Organization* 55 (03): 693–715.

Kiras, James D. 2013. “Irregular Warfare: Terrorism and Insurgency”. In *Strategy in the Contemporary World: an Introduction to Strategic Studies,* edited by John Baylis, James J Wirtz and Colin S Gray. Oxford: Oxford University Press.

Kramer, Mark. 2005. “Guerrilla Warfare, Counterinsurgency and Terrorism in the North Caucasus: The Military Dimension of the Russian-Chechen Conflict”. *Europe-asia Studies* 57 (2): 209–290.

Kreutz, Joakim. 2010. “How and When Armed Conflicts End: Introducing the UCDP Conflict Termination Dataset”. *Journal of Peace Research* 47 (2): 243–250.

Ladwig, Walter C. 2015. “Indian Military Modernization and Conventional Deterrence in South Asia”. *Journal of Strategic Studies* 38 (5): 729–772.

Larson, Eric V, and John E Peters. 2001. *Preparing the U.S. Army for Homeland Security: Concepts, Issues, and Options.* Santa Monica: Rand Corporation.

Lee, Peter M. 2012. *Bayesian Statistics: An Introduction.* Sussex: Wiley.

Lee, Terence. 2009. “The Armed Forces and Transitions from Authoritarian Rule: Explaining the Role of the Military in 1986 Philippines and 1998 Indonesia”. *Comparative Political Studies* 42 (5): 640–669.

Lipset, Seymour Martin, and Stein Rokkan. 1967. “Cleavage Structures, Party Systems, and Voter Alignments: An Introduction”. In *Party Systems and Voter Alignments,* 1–64. New York: Free Press.

Little, Roderick J. A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data.* 2nd. New Jersey: Wiley.

Lutscher, Philipp M. 2016. “The More Fragmented the Better?The Impact of Armed Forces Structure on Defection during Nonviolent Popular Uprisings”. *International Interactions* 42 (2): 350–375.

Lyall, Jason. 2009. “Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya”. *Journal of Conflict Resolution* 53 (3): 331–362.

Lyall, Jason, and Isaiah Wilson. 2009. “Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars”. *International Organization* 63 (1): 67–106.

Mandeville, Bernard. 1732. *The Fable of the Bees, or Private Vices Publick Benefits.* Oxford: Clarendon Press.

Mason, David T. 2004. *Caught in the Crossfire: Revolution, Repression, and the Rational Peasant.* Oxford: Rowman & Littlefield Publishers, Inc.

McElreath, Richard. 2015. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.* Boca Raton: CRC Press.

McLauchlin, Theodore. 2015. “Desertion and Collective Action in Civil Wars”. *International Studies Quarterly* 59 (4): 669–679.

McNeish, Daniel. 2016. “On Using Bayesian Methods to Address Small Sample Problems”. *Structural Equation Modeling: A Multidisciplinary Journal* 23 (5): 750–773.

Melander, Erik, Therése Pettersson and Lotta Themnér. 2016. “Organized Violence, 1989\textendash{}2015”. *Journal of Peace Research* 53 (5): 727–742.

Millett, Allan R, Williamson Murray and Kenneth H Watman. 1987. *The effectiveness of military organizations.* Washington, D.C.: US Department of Defense.

Nepstad, Sharon Erickson. 2011. *Nonviolent Revolutions: Civil Resistance in the late 20th Century.* Oxford: Oxford University Press.

———. 2013. “Mutiny and nonviolence in the Arab Spring”. *Journal of Peace Research* 50 (3): 337–349.

Pachon, Alejandro. 2014. “Loyalty and Defection: Misunderstanding Civil-Military Relations in Tunisia During the ‘Arab Spring’”. *Journal of Strategic Studies* 37 (4): 508–531.

Peaty, John. 1999. “Haig and Military Discipline”. In *Haig: a Reappraisal 70 Years On,* edited by Brian Bond and Nigel Cave, 271. Pen & Sword Books, Ltd.

———. 2002. “The Desertion Crisis in Italy, 1944”. *Royal United Services Institute (RUSI)* 147 (3): 76–83.

Petersohn, Ulrich. 2017. “Private Military and Security Companies (PMSCs), Military Effectiveness, and Conflict Severity in Weak States, 1990–2007”. *Journal of Conflict Resolution* 61 (5): 1046–1072.

Pillalamarri, Akhilesh. 2014. *India’s Anti-Sikh Riots, 30 Years On.* The Diplomat. *URL: https://thediplomat.com/2014/10/indias-anti-sikh-riots-30-years-on/, Retrieved on: 21.12.2017.*

Pilster, Ulrich, and Tobias Böhmelt. 2011. “Coup-Proofing and Military Effectiveness in Interstate Wars, 1967-99”. *Conflict Management and Peace Science* 28 (4): 331–350.

———. 2012. “Do Democracies Engage Less in Coup-Proofing? On the Relationship between Regime Type and Civil–Military Relations”. *Foreign Policy Analysis* 8 (4): 355–372.

Pilster, Ulrich, Tobias Böhmelt and Atsushi Tago. 2016. “The Differentiation of Security Forces and the Onset of Genocidal Violence”. *Armed Forces & Society* 42 (1): 26–50.

Pion-Berlin, David, Diego Esparza and Kevin Grisham. 2014. “Staying Quartered”. *Comparative Political Studies* 47 (2): 230–259.

Powell, Jonathan. 2019. “Leader Survival Strategies and the Onset of Civil Conflict: A Coup-Proofing Paradox”. *Armed Forces & Society* 45 (1): 27–44.

Powell, Jonathan M. 2012a. *Coups and Conflict: the Paradox of Coup-Proofing.* University of Kentucky: PhD Dissertation.

———. 2012b. “Determinants of the Attempting and Outcome of Coups d’état”. *Journal of Conflict Resolution* 56 (6): 1017–1040.

Quinlivan, James T. 1999. “Coup-Proofing: Its Practice and Consequences in the Middle East”. *International Security* 24 (2): 131–165.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. “Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature”. *Journal of Peace Research* 47 (5): 651–660.

Ray, Jayanta Kumar. 2007. *Aspects of India’s International Relations, 1700 to 2000: South Asia and the World.* Kolkata: Pearson Education India.

Roessler, Philip G. 2016. *Ethnic politics and state power in Africa: the logic of the coup-civil war trap.* Cambridge: Cambridge University Press.

Rose, Elihu. 1982. “The Anatomy of Mutiny”. *Armed Forces & Society* 8 (4): 561–574.

Ruffa, Chiara. 2014. “What Peacekeepers Think and Do: An Exploratory Study of French, Ghanaian, Italian, and South Korean Armies in the United Nations Interim Force in Lebanon”. *Armed Forces & Society* 40 (2): 199–225.

Ruggeri, Andrea, Han Dorussen and Theodora Ismene Gizelis. 2017. “Winning the Peace Locally: Un Peacekeeping and Local Conflict”. *International Organization* 71 (1): 163–185.

———. 2018. “On the Frontline Every Day? Subnational Deployment of United Nations Peacekeepers”. *British Journal of Political Science* 48 (4): 1005–1025.

Schiel, Rebecca, Christopher Faulkner and Jonathan Powell. 2017. “Mutiny in Côte d’Ivoire”. *Africa Spectrum* 52 (2): 103–115.

Schutte, Sebastian. 2017. “Geographic determinants of indiscriminate violence in civil wars”. *Conflict Management and Peace Science* 34 (4): 380–405.

Silva, Romesh, Jasmine Marwaha and Jeff Klingner. 2009. *Violent Deaths and Enforced Disappearances During the Counterinsurgency in Punjab, India: A Preliminary Quantitative Analysis.* A Joint Report by Benetech’s Human Rights Data Analysis Group / Ensaaf, Inc. *URL: http:// www.ensaaf.org/publications/reports/de Retrieved on: 22.12.2017.*

Singh, Gurharpal. 2000. *Ethnic Conflict in India: A Case-Study of Punjab.* London: Macmillan Press.

Singh, Sangat. 2002. *The Sikhs in History.* New York: Singh Brothers.

SIPRI. 2017. *SIPRI Military Expenditure Database.* Stockholm International Peace Research Institute.

Smith, Adam. 1976. *An Inquiry into the Nature and Causes of the Wealth of Nations [reprint].* The University of Chicago: The University of Chicago Press.

Soeters, Joseph. 2018. “Organizational cultures in the military”. In *Handbook of the sociology of the military,* edited by G. Caforio and M. Nuciari, 251–272. Springer.

Soeters, Joseph L., Cristina-Rodica Poponete and Joseph T. Page Jr. 2006. “Culture’s Consequences in the Military”. In *Military Culture, Military life: The psychology of serving in peace and combat: Military culture,* edited by T. W. Britt, A. B. Adler and C. A. Castro. Prague: Praeger Security International.

Spear, Joanna. 2008. “Counterinsurgency”. In *Security Studies: an Introduction,* edited by Paul D Williams. London: Routledge.

Spolaore, Enrico, and Romain Wacziarg. 2016. “Ancestry, Language and Culture”. In *The Palgrave Handbook of Economics and Language,* 174–211. London: Palgrave Macmillan UK.

Staniland, Paul. 2008. “Explaining Civil-Military Relations in Complex Political Environments: India and Pakistan in Comparative Perspective”. *Security Studies* 17 (2): 322–362.

———. 2017. “Armed politics and the study of intrastate conflict”. *Journal of Peace Research* 54 (4): 459–467.

Stephan, Maria J., and Erica Chenoweth. 2008. “Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict”. *International Security* 33 (1): 7–44.

Sudduth, Jun Koga. 2017. “Coup risk, coup-proofing and leader survival”. *Journal of Peace Research* 54 (1): 3–15.

Sullivan, Patricia L. and Karreth, Johannes. 2015. “The conditional impact of military intervention on internal armed conflict outcomes”. *Conflict Management and Peace Science* 32 (3): 269–288.

Sundberg, Ralph, and Erik Melander. 2013. “Introducing the UCDP Georeferenced Event Dataset”. *Journal of Peace Research* 50 (4): 523–532.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule.* Cambridge: Cambridge University Press.

———. 2013. “Contracting on Violence”. *Journal of Conflict Resolution* 57 (5): 765–794.

Talmadge, Caitlin. 2015. *The Dictator’s Army: Battlefield Effectiveness in Authoritarian Regimes.* London: Cornell University Press.

Telford, Hamish. 2001. “Counter-Insurgency in India: Observations from Punjab and Kashmir”. *Journal of Conflict Studies* 21 (1): 1–29.

Thompson, William R. 1975. “Regime Vulnerability and the Military Coup”. *Comparative Politics* 7 (4): 459–487.

———. 1976. “Organizational Cohesion and Military Coup Outcomes”. *Comparative Political Studies* 9 (3): 255–276.

Thornton, Rod. 2000. “The role of peace support operations doctrine in the British Army”. *International Peacekeeping* 7 (2): 41–62.

Tollefsen, Andreas Forø, Håvard Strand and Halvard Buhaug. 2012. “PRIO-GRID: A unified spatial data structure”. *Journal of Peace Research* 49 (2): 363–374.

Van Creveld, Martin L. 2005. *Kampfkraft: Militärische Organisation und Leistung der deutschen und amerikanischen Armee 1939-1945.* Graz: Ares.

Welch, Claude Emerson. 1976. *Civilian Control of the Military: Theory and Cases from Developing Countries.* New York: New York Press.

Wickham-Crowley, Timothy P. 1992. *Guerrillas and Revolution in Latin America: A Comparative Study of Insurgents and Regimes Since 1956.* Princeton: Princeton University Press.

Williams, Paul D. 2008. “War”. In *Security {{Studies}}: An {{Introduction}},* edited by Paul D Williams. London: Routledge.

Wood, Reed M., and Jacob D. Kathman. 2014. “Too Much of a Bad Thing? Civilian Victimization and Bargaining in Civil War”. *British Journal of Political Science* 44 (03): 685–706.

Zhukov, Yuri M. 2015. “Population Resettlement in War: Theory and Evidence from Soviet Archives”. *Journal of Conflict Resolution* 59 (7): 1155–1185.

University of Essex

UNIVERSITY OF ESSEX
POSTGRADUATE RESEARCH EDUCATION TEAM

Statement of Authorship

Thesis Title	Security forces in conflict: Composition, specialisation, and effectiveness in counterinsurgency and peacekeeping operations	
Publication Status	☐ Published Work	☐ Accepted for Publication
	☒ Submitted for Publication	☐ Unpublished and Un-submitted work written in manuscript style, of publishable quality
Publication Details	Submitted for consideration to Paper 1 Force structure and local peacekeeping effectiveness: Micro-level evidence from Sub-Saharan Africa.	

Principal Author

Name of Principal Author (Candidate)	Christoph Dworschak		
Registration Number	1700127		
Department	~~Choose an item.~~ Government		
Contribution to the Paper	Abstract, literature, theory, empirics		
Overall percentage (%)	55		
Signature		**Date:**	01/06/2020 (DD/MM/YY)

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate to include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100%, less the candidate's stated contribution.

Name of Co-Author	Deniz Cil
Contribution to the Paper	Abstract, literature, theory, empirics

Overall percentage (%)	45		
Signature		**Date:**	02/06/2020 (DD/MM/YY)

Name of Co-Author			
Contribution to the Paper			
Overall percentage (%)			
Signature		**Date:**	(DD/MM/YY)

Please copy and paste additional co-author panels here as required.