Non-Unitary Actors in Intrastate Armed Conflict
A Quantitative Analysis of Civilian Defense Forces’ Effect on Intrastate War Duration

Bachelor Thesis By:
AnnaVera Eriksson
Department of Peace and Conflict
January 2018

Supervisor: Sabine Otto
“There is no instance of a nation benefitting from prolonged warfare”

- Sun Tzu
# Table of Contents

1. Introduction........................................................................................................................................4
   1.2. Previous Research .........................................................................................................................6
2. Theory ................................................................................................................................................10
   2.1. Central Concepts To The Research ...............................................................................................10
   2.2. The Theoretical Argument ............................................................................................................11
3. Research Design ................................................................................................................................15
   3.1. Method ..........................................................................................................................................15
   3.2. Scope Conditions & Limitations .....................................................................................................15
   3.3. Dependent Variable – Duration ......................................................................................................16
   3.4. Independent Variable – Presence Of CDF .....................................................................................17
   3.5. Control Variables ..........................................................................................................................18
   3.6. Validity & Reliability .....................................................................................................................20
   3.7. Data & Source Criticism .................................................................................................................22
4. Results & Analysis .................................................................................................................................24
   4.1. Descriptive Statistics ......................................................................................................................24
   4.2. Results ..........................................................................................................................................26
   4.3. Control For Missing Values ...........................................................................................................28
   4.4. Robustness Checks .........................................................................................................................29
   4.5. Discussion .....................................................................................................................................32
5. Summary & Conclusion .........................................................................................................................35
6. References .........................................................................................................................................36
Appendix ..................................................................................................................................................40
1. Introduction

Since the beginning of the 21st century the number of conflicts in the world have decreased. In spite of this encouraging trend, data also tells us that the conflicts of today have grown more violent, resulting in an overall increase in conflict related fatalities (Norton-Taylor, 2015). According to data from the World Bank (2017), there are currently two billion people living in countries affected by conflicts, fragility or violence and 80 percent of all the humanitarian need in the world are due to conflicts. Conflicts have a dire effect on a large portion of the world we live in. Understanding the dynamics which drive conflict is imperative for conflict management and prevention.

Grasping why conflicts endure is a key aspect for understanding how to reach conflict resolution. If we can understand how to break the patterns of why conflicts continue, much human suffering can be avoided. Therefore, the principal source of motivation where this thesis finds its research question relates to understanding the conflict dynamics which drive duration.

*Why do some intrastate armed conflicts last longer than others?*

Many scholars have studied the causes for conflict duration. It has been found that several attributes which affect the duration of conflict relate to the causes of conflict onset (Fearon, 2004). Thus, the reasons for civil war onset can be key aspects to why some conflicts are easily resolved while others are not. Consequently, conflict onset theories have been a stepping stone for scholars within the field of conflict duration.

Well known and established theories to explain conflict onset are greed and grievance. Greed relates to the possibility of profits which motivate rebels to initiate and to continue armed conflicts. The basis of the theory claims that if there is an opportunity to gain something from armed conflict, an armed conflict will commence and endure (de Soysa 2000; Collier and Hoeffler, 2004). Grievance relates to socio-economic, ethnic or other inequalities which may cause some groups in society enough suffering to take up arms. This type of onset would mean that a conflict endures as long as the incompatibility is unsettled (Collier, Hoeffler, and Söderbom 2004: DeRouen and Sobeck 2004).

Apart from the traditional explanations, Cunningham et al (2009:575) present an interesting new aspect of conflict dynamics. The scholars argue in their study that the rebel actor should be disaggregated to lower levels of analysis in order to properly understand the actor’s effect on the conflict dynamics. The scholars claim that rebel group’s offensive and defensive capabilities play different roles in relation to the conflict’s duration.
Studying rebels as non-unitary actors is thought-provoking and raise questions about whether the government should be studied as unitary. Studying the government as a non-unitary actor, and how this can affect conflict, can be found within research on Pro Government Militias (PGMs). It has been found that using certain types of militias can have a substantial effect on the probability of government victory in counterinsurgent wars (Peic, 2014). Government deployment of Civilian Defense Forces (CDFs) increase the probability of government victory in conflict (Peic, 2014: 174). If the government actor’s complexity affects outcome, and it has been found that the disaggregation of the rebel actor contributes to the analysis of conflict duration, the constellation of the government actor’s effect on duration becomes a natural question.

Thus far, findings on PGMs show that different types of PGMs combined with different types of government, has an effect on conflict duration. Strong states using PGMs might shorten a conflict while autonomous PGMs used by weak states might lengthen it (Jentzsch, Kalyvas, and Schubiger, 2015). This research recommends further investigation, stating a research gap concerning the determinants of different types of militias’ effects on duration. Addressing this research gap might render new evidence for studying effects on civil war duration.

Therefore, building on previous research on conflict duration as well as the developing field on Pro-Government Militias (PGMs), this thesis aims to explore the effects a certain constellation of the government actor might have on conflict duration. This study claims that using Civilian Defense Forces (CDFs) in intrastate armed conflict has a negative effect on conflict duration. This is due to the change in the information asymmetry between the conflicting parties. Because of their proximity to the civilian support base of the rebels, CDFs are able to provide the government actor with information otherwise concealed from them. With a stable source of information about the rebels, the circumstances change and bargaining space can be found, leading to more quick conflict resolution.

The effects Civilian Defense Forces have on conflict duration is explored through a quantitative comparative analysis of civil armed conflicts between 1981-2007. As theoretical framework, rational choice theories of war are used to understand why, when governments deploy CDFs, the conflict dynamics are altered which in turn has consequences for the conflict’s duration. The findings suggest that CDFs substantially increase the duration of civil armed conflict. However, these findings do not necessarily reject the presented theory as methodological limitations and possible endogeneity opens avenues for future research diving deeper into the subject.

The thesis begins with a summary of previous research, a discussion of the proposed theory followed by an outline of the research design. Thereafter, a presentation of the results leads to
analysis. This is followed by a discussion of the findings and proposals for future research. Finally, a summary and conclusion will be presented.

1.2. Previous Research

1.2.1. Duration of Civil Armed Conflict

Duration of civil armed conflict is a well-studied field approached with an abundance of methods. Generally, studies are conducted with large samples attempting to find generalizable answers as to what triggers conflicts to endure.

Fearon (2004) conducts an inductive analysis of civil wars in order to find patterns in causes for duration. Using a multivariate analysis and constructing a game model the study explains the factors which cause the variance in duration of civil wars. He finds five factors which noticeably affect duration. Civil wars which are started near the center of power tend to be shorter, explaining why civil wars triggered by coups tend to end more quickly. Thus, conflicts that have their locus in the peripheries of a nation tend to last longer even though rebels of this kind most commonly are substantially weaker in military power than the state forces (2004:277). The study also finds strong empirical patterns which indicate that conflicts which involve ethnic minorities in the peripheries of a nation are substantially longer than other wars. This issue, often referred to as “sons of the soil” causes markedly longer conflicts (ibid, 2004:283). Another factor with stark importance on duration is the presence of contraband. In conflicts where rebels are able to finance themselves, or partially finance themselves by smuggling products such as diamonds, cocaine, opium or other commodities, conflicts are found to endure (ibid, 2004:278). Lastly, he finds patterns indicating that civil wars in Eastern Europe as well as anti-colonial wars tend to be shorter (ibid, 2004:280-282).

Duration has also been found to depend on the state’s capacity. Analyzing the strength of state armies, it has been found that states using a strong army towards an insurgency only are effective if the state is capable of defeating the insurgents swiftly (DeRouen and Sobeck, 2004). If not, the strong army will only exacerbate the conflict and contribute to its duration (DeRouen and Sobeck 2004:317). Furthermore, DeRouen and Sobeck (2004:314) found that the geographic qualities of a state can affect the duration of a conflict within the state; forests increase the duration of conflict while mountainous terrain merely affects the probability of either party’s victory.

Concerning either parties victory, Cunningham et al (2009) also find that governments generally are more likely to win if they are able to defeat rebels close to conflict onset. Moreover, strong rebels tend to fight short wars. The duration of a civil war also seem to increase when the rebels are weaker than the government but are in control of territory in the periphery of the state.
(Cunningham et al, 2009:573). This means that the rebels’ capabilities need to be understood in offensive capabilities (being able to inflict cost on the government) and defensive capabilities (being able to withdraw and evade government repression) (ibid, 2009:575). The study recommends a reevaluation of the importance of military power as the answer to insurgencies and recommend that nonviolent dispute resolution might be preferable if governments are unable to defeat insurgencies at an early stage of conflict (ibid, 2009:593).

The outcome and duration of intrastate armed conflict has been found to change substantially when more than two players are involved in said conflict. Studying the effects of veto players in civil war, Cunningham (2006) claim that conflicts which involve more than two actors are harder to resolve and therefore continue longer. The study uses a multiparty bargaining model drawn from theories in political bargaining to examine the effects multiple parties have on intrastate armed conflicts. Claiming that the bargaining range is altered when multiple parties enter into conflict or negotiation, the article explains that the room for overcoming commitment and information problems, which should be possible to overcome within bargaining range, grows smaller with more than two parties in a conflict. This is an especially important issue in multiparty intrastate armed conflicts as information asymmetric are greater in civil wars (ibid, 2006:879).

It has also been found that not only the primary parties of a conflict might affect its duration. Involving third parties as interveners is found to affect both duration and outcome. Regan (2002:59) state that third party intervention alters the estimated chances of victory for the warring parties by changing the balance of capabilities between them. Intervening on either side of the conflict will have an immediate effect on the conflict and a disproportionately bigger effect on the rebel side as the battlefield results have more important implications on their recruitment. Furthermore, the timing of the intervention will affect the parties willingness to settle now rather than later (Regan, 2002:60). The study claims that intervention on the government side should have a shortening effect on the conflict. Intervention of the rebel side will have a parity effect on the balance of capabilities, moving the parties further away from a hurting stalemate and creating longer conflicts. However, the results of the research show that any third-party intervention has a positive effect on the duration of a conflict, and no statistical significance is found on the intervention of the government side and shorter conflicts. Conflict which have interventions supporting both sides have a considerably greater chance of continuing than do conflicts which do not have interventions supporting both sides. Thus, neutral interventions have a greater probability of continuing a conflict than conflicts with biased interventions (ibid, 2002:71-72).

Generally, most studies are based on the assumption that the government can be considered a unitary actor. What happens if the government is not a unitary actor? A field which has
disaggregated the state actor is the field of Pro Government Militias which examines the change in societal dynamics when the government renounce their monopoly on violence.

1.2.2. Pro-Government Militias (PGMs)

Research on PGMs has previously not delved deeply into analysis of the affects militias might have on duration. However, one study conducted by Jentzsch, Kalyvas, and Schubiger (2015) explores the subject. The study (ibid, 2015) claims that the type of PGM as well as the function of the militia can affect conflict duration. Strong states using militias might shorten the length of conflict while more autonomous PGMs used by weaker states might lengthen it. Furthermore, PGMs have been found to play a role in the type of violence used in civil wars. Governments have been thought to use the possibility to outsource violence they do not want to be connected to through using PGMs. Stanton (2015) explores the phenomena of violence against civilians by comparing the actions of PGMs in relation to regular armed forces. It is found that militias engage in violence against civilians when the government’s official troops do so, but they do not to any large extent engage in these actions when governments restrain from it. Thus, correspondence between militia and government behavior has been found (ibid, 2015).

During conflict, a specific type of PGMs (CDFs) has been found to increase the governments access to intelligence, making counterinsurgency campaigns more efficient (Peic, 2014). However, during these types of conflict, the advantage PGMs pose for the government has also been found to increase the overall fatalities of the conflicts (Clayton and Thomson, 2016). Peic (2014) argue that governments use Civilian Defense Forces (CDFs), in order to reestablish credibility and provide security in counterinsurgency wars. With CDFs, that are local and recruited from the civilian population, governments can show presence and counterbalance some of the inherent disadvantages insurgencies pose for the governments when rebels use irregular tactics. When using CDFs, governments can obtain more accurate and timely intelligence on insurgents which gives them the possibility of punishing civilians selectively instead of indiscriminately. As civilians are the primary support base of insurgents, when governments are able to obtain civilian support through recruiting civilian key players, insurgents are weakened and experience difficulties in employing guerilla tactics. Governments will then be able to reap further public support as a result of successful counterinsurgency campaigns. Thus, in counterinsurgency wars where governments utilize CDFs, the probability of government victory is increased (Peic 2014).

To complement the findings of the usefulness of CDFs for governments, Clayton and Thomson (2016) have studied the effects using militias have on the levels of violence in intrastate conflicts. When governments deploy CDFs, they are increasingly able to correctly target insurgents
and avoid indiscriminate violence. However, as they turn civilians who might formerly have supported rebels against rebels, two mechanisms are triggered which both increase the violence in conflicts. When civilians change their backing, indiscriminate violence from rebels towards civilians is increased. The violence serves as a threat towards defecting as well as revenge towards those who have already defected. The violence within insurgencies is also found to increase since defectors cause splintering within the groups. In total, the conflict is found to be more violent with CDFs present. Important to emphasize however, it that these findings are not claimed to be generalizable to all forms of PGMs (Clayton and Thomson, 2016).

Carey et al. (2015) also acknowledge that there are intelligence advantages for governments in using militias in counterinsurgencies. However, primarily, Carey et al (2015) focus on the political advantages produced by using PGMs. They identify that in weak democracies with high reliability on foreign aid, PGMs are likely used by governments in order to evade accountability of their violent actions as they are able to plausibly deny their involvement in the repression (ibid, 2015). The study also finds that there are logistic incentives, such as the increase in force numbers without the increase in expenses, which motivates the utilization of PGMs.

The risks for the government of using PGMs is analyzed by Carey and Mitchell (2017). They find that the type of members in a PGM affect the risk it poses towards the government. Local or non-civilian militias pose little threat towards the stability to the state compared to those groups that are capable of recruiting larger parts of the society (Carey and Mitchell, 2017).

It can be concluded that even though previous research has not entirely overlooked the possible impact of PGMs on conflict duration, the field has generally focused elsewhere. As presented in this review, scholars have recommended to disaggregate the government actor in order to understand the dynamics that drive duration. The PGM research field presents this possibility. Therefore, this thesis will combine these fields in order to examine what effects CDFs have on conflict duration.
2. Theory

2.1. Central Concepts to the Research

2.1.1. Pro Government Militias & Civilian Defense Forces

Research has shown that governments rarely act as a unitary actor monopolizing the violence of the state. Data from Carey, Mitchell and Lowe (2013) show that Pro-Government Militias occur in 81 percent of all country-year intrastate conflicts between 1981 and 2007. A PGM is an armed group linked to the government informally or officially but exist outside of the regular government security apparatus, and has some level of organization (Carey and Mitchell 2017: 128). This specific type of armed groups has been conceptualized in order to understand the causes and consequences of governments’ outsourcing violence (ibid, 2017:128). PGM is a quite thick concept including all types of government affiliated armed groups, be they large armed organizations or small local groups, well-equipped and quasi-professional or ad hoc formed and poorly equipped with machetes and bows (Jentzsch, Kalyvas, and Schubiger, 2015:177-178).

Civilian Defense Forces (CDFs) is a subcategory of this phenomena. Generally, and for the purpose of this thesis, a CDF is defined as an armed group of predominantly non-combatants, recruited or created by the government. They are also distinguished through their local connection to the area wherein they operate which is located in the area of conflict. The group should also be ethnically affiliated to the rebel actor if the conflict has ethnic dimensions. Usually, the CDF is immobile and the members of the force remain in their own homes rather than organize themselves in a specific location and operate in a larger area (Peic 2014:165; Clayton and Thomson, 2016:500). CDFs can be created by the state with the objective to show governmental presence where they do not have the capacity to deploy regular troops, or in order to collect intelligence from the civilian population and sway popular opinion. CDFs can also be created out of grass root initiatives to later be recruited by the state apparatus (Peic 2014:165-166).

2.1.2. Internal Armed Conflict & Duration

What constitutes an armed conflict is a well-discussed subject creating different definitions and operationalizations depending on the scholar’s research goals and use of data. In this study, a definition using a low threshold in terms of battle-related deaths for armed conflict is used in order to capture a greater variance in the data. This thesis will utilize the definition created by the Uppsala Conflict Data Program (UCDP) where a conflict is “a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25-battle related deaths in one calendar year”
The terms civil war, intrastate armed conflict and internal armed conflict are in this thesis used interchangeably and all refer to this definition.

It is problematic to capture the entire concept of conflict or war in a small number of reference points since the thickness of the concept grows greater the more it is scrutinized. Therefore, the definition should, beyond the number of deaths, at least also take into account the subject of the dispute as well as the presence of opposing parties. Even though the definition used here captures these parts of the concept, this simplification of the concept does create possible issues of validity. However, the high reliability of the dataset relating to this definition, and the usability of it in relation to the objective of this thesis, well compensates for issues in validity. A great asset in this conceptualization lies within the clear definition of when a conflict starts and when it ends. Through this, duration is easy to understand and analyze as the time elapsed between the start and end date of the conflict.

2.2. The Theoretical Argument

In order to theorize on why some conflicts endure, one must understand why they are initiated and how they are resolved. It is through the theories of conflict onset and resolution this thesis builds its theoretical framework. This section will present and discuss the theory of information asymmetry which is the core of the theoretical argument of this thesis. In order to understand the dynamics and implications of this theory, the foundation for it which is found in rational choice theory and bargaining space, will be presented initially. Thereafter, the aspects of information asymmetry will be presented and discussed.

A theoretical framework commonly used by scholars in conflict duration revolve around theories of cost/benefit analysis and balance of capabilities (Balcells and Kalyvas 2014; Cunningham 2006; DeRouen and Sobeck 2004). The balance of military capabilities together with how the parties interact with each other (their military symmetry) is claimed by many scholars to be the fundamental reason for how the duration and intensity of conflicts develop (Balcells and Kalyvas, 2014:1393). Cost/benefit calculation and balance of capabilities are theories derived from rational choice theories of conflict. Rational choice theories assume that all actors are rational and act according to continuous calculations of the most beneficial outcome for themselves (Fearon 1995:381).

To explain the most beneficial outcome, bargaining space is a theoretical concept used to describe the actors' tug-of-war over incompatibilities. The theory explains that there is a constant space between the actors' highest accepted cost and lowest accepted benefits wherein both can be satisfied. However, due to insufficient information, parties can choose conflict even though there
is a theoretical space for them both to be satisfied without the cost of violence. The possible bargaining space is present throughout the entire conflict. Yet, since they only learn more about each other’s capabilities and motivations throughout the conflict, the amount of the bargaining space each party can claim for themselves change in relation to the events that unfold in the conflict (Collier, et al., 2004:254).

Consequently, information asymmetry can be a cause for conflict onset as the parties are assumed to have little information about the opponents’ capabilities and motivations during the beginning of a conflict (Walter 2009:245). In intrastate wars, governments struggle in understanding how strong a potential rebel group is, both in concern to their military power but also how powerful their motivation and resolve is. It is assumed that governments only concede to rebel groups which are powerful enough to inflict painful costs. However, governments generally do not possess enough information to correctly evaluate the potential rebels relative strength before armed conflict onset (Walter 2009:246). It is also assumed that the government always is in an information disadvantage in civil war as the rebels more easily can hide information from the government than vice versa. However, the rebels do not have complete information about the government either. Thus, private information creates a situation where rebel and government actors act without complete knowledge of the situation and therefore may choose conflict. Bargaining theory explain that actors learn more about the opponents’ private information via the violent acts committed in the conflict. Information can be deducted through actions since the actions committed in the conflict are assumed to mirror what the actor has the capability to do, and what the actor is willing to do in relation to the costs the action might cause (Zartman, 2000:228).

As the conflict evolves, actors eventually learn enough about each other to either defeat one another, or enter into a settlement as they realize the most beneficial option. The reasoning then follows, that the duration of a conflict is also determined by the information they have. With information, the actors can make continuous cost/benefit calculations in relation to how they perceive the balance of capabilities and the opponent’s motivation. Hence, actors endure in a war if they believe that they have more to gain from fighting than from yielding or agreeing to a settlement (Fearon 1995:386). They will act rationally according to the information they have, not necessarily according to the actual situation.

As stated above, all information is not available in conflicts. This is an even more acute issue in intrastate conflicts making the issue relevant as a mechanism. Information asymmetries are likely to be more common and more severe in intrastate conflicts than in interstate conflicts because the government is less likely to be able to obtain information about the opponents capabilities (Walter
In intrastate conflict, the rebels might also not know their own capabilities before they engage in battle as they are assumed to earn the majority of their support once the conflict is commenced. Even so, in those cases where the rebels know their own strength, there are strong incentives to withhold this knowledge from the government since the government can use it to repress the rebels (ibid, 2009:246). Consequently, a civil conflict is more likely to occur, and endure, where information about capabilities is most uncertain. This uncertainty can be due to how non-obvious the rebel’s capabilities are, or how hard it is for governments to monitor the rebel’s strength and support (ibid, 2009:250).

When Civilian Defense Forces (CDFs) are deployed in a conflict, the information asymmetry changes (Peic, 2014:163). In intrastate armed conflicts where rebel tactics are utilized, one of the most challenging obstacles for the government to overcome relates to the access to information about the opponents who hide amongst the civilians they are supported by (ibid, 2014:162). Without detailed knowledge about the opponents, the government is more prone to use indiscriminate violence against the civilian population in order to target the rebels (Clayton and Thomson, 2016). When CDFs are created, the government gains access to a new source of information within the civilian population. This stable source of information gives the government more accurate knowledge about where the rebels are active or hiding. As mentioned in the definition, CDFs are civilians recruited because they live in the communities which serve as the support base for rebels. As a part of the community, the members of CDFs can interact with rebels and their supporters in a manner not possible for government officials or their regular proxies. The proximity of the CDF to the rebels gives them information which they can transmit to the government, who otherwise would struggle to attain that information. Consequently, members of CDFs can give private information to government forces about rebel capabilities and motivations. In other words, formerly private information is revealed through the use of CDFs, information asymmetries are reduced and bargaining space is exposed.

The use of CDFs might then lead to the government gaining an advantage in the conflict, using the information to strike at rebels in ways which leads to their defeat. As stated in the literature review, the findings of Peic (2014) show that governments more often win in counterinsurgencies when using CDFs. This is because they are able to target rebels more accurately when using CDF-provided information. Yet, the deployment of CDFs could arguably also lead to settlement. The information procured which reduces information asymmetry could reveal information about the bargaining space which shows the government that they are inferior to the rebels. CDF-generated information could then lead to settlement. Thus, even though the provided information can help governments to win, this is most probable if the balance of
capabilities is in favor of the government. Otherwise, it is possible that the new information instead leads to settlement. Thus, the presence of CDFs alleviates information asymmetries and helps the parties find bargaining space which otherwise only would be found through continued fighting (Zartman 2000:228). From this reasoning, the hypothesis of this thesis is drawn.

\[ H1: \text{Civil wars in which governments deploy CDFs are shorter than wars without CDFs.} \]

Nonetheless, there is of course a third possibility in the outcome of civil war where rebels become victors. To this author’s knowledge however, most studies find that irregular warfare most often result in government victory, even though duration might be affected (Balcells and Kalyvas, 2014; Peic, 2014; Clayton and Thomson, 2016). Even so, it is possible that the presence of CDFs increase confusion in the conflict contributing to duration. With the use of militias, the government is no longer a unitary actor in the eyes of the rebels and the balance of capabilities becomes unclear for the conflicting parties. Confusion leads to an increase in the information asymmetry which then in turns necessitates more combat interaction for the actors to realize the bargaining space. When insurgent conflicts endure, the rebel actor increases its chance of winning (DeRouen and Sobeck, 2004) possibly diminishing the effect of the CDF, and explaining how rebels might still be able to win conflicts where CDFs are present.

However, it is also possible that the added CDF indicate a shift in the balance of capabilities which pressures the rebels to use increased force as they fear being outmaneuvered. This could then escalate the intensity of the conflict which is indicated in the findings of Clayton and Thomson (2016) where CDFs are found to cause increased rebel violence. The result could then be a victorious outcome for the rebels, given that the government is weaker than calculated and that the rebels are able to utilize guerilla tactics effectively after surviving the initial intensification of violence from the government side (DeRouen and Sobeck, 2004). However, regardless of the conflict outcome, the deployment of CDFs should function as a catalyst for the conflict, eventually leading to a swifter conflict termination than in conflicts without the presence of CDFs, still giving support to the presented hypothesis.

Other research has acclaimed militias to serve as a force multiplier changing the balance of physical capabilities in the governments favor (Carey, et al., 2015: 851). The militia’s function would then relate to the actual fighting and the government’s ability to lower costs, rather than the release of private information affecting the bargaining space. This might be true for Pro-Government Militias as an aggregated unit of analysis. militias with large bodies of members, relatively strong organizations and with access to relatively advanced equipment have been found to increase the
government’s military capabilities (Carey and Mitchell 2017:136). Furthermore, some militias are mobile and work semiofficially or informally in larger geographical areas as an extension of state troops (ibid, 2017:130). However, this theory is not true for Civilian Defense Forces since they are conceptualized as local, civilian groups recruited for defensive purposes (Clayton and Thomson 2016:502). The contribution CDFs make to the military capabilities of the government actor are simply too meager to plausibly change the balance of capabilities between the opposing parties, explaining why this thesis does not use the balance of capabilities as a theoretical framework. CDFs are civilians, they do not contribute substantially to the actual fighting. It is thus reasonable to assume that the effect they contribute with concerns other aspects of the conflict. If CDFs have a negative effect on duration, the most plausible explanation to this effect lies within their access to private information.

3. Research Design

3.1. Method

The aim of this thesis is to present and analyze a theory on the effects Civilian Defense Forces have on intrastate conflict duration. The most common method used to study duration is by using the Cox Proportional Hazard Model (Cunningham, 2006:882), however, for the purpose of this study, the theory is tested using Ordinary Least Squared (OLS) regression as quantitative method. The unit of analysis is duration of internal armed conflicts in months. The theory is tested on observations between 1981 and 2007 building on the CDF data coded and presented by Clayton and Thomson (2016). This section outlines the methodology of the analysis, the variables and their operationalizations, the scope conditions and limitations as well as a discussion of reliability and validity of the measurements.

3.2. Scope Conditions & Limitations

The limitations of this analysis are determined by the available data on CDFs and their occurrence in conflicts. The Pro-Government Militia Database (PGMD) is the maverick database on the occurrence of different types of PGMs in the world between 1981-2007 (Carey, Mitchell, and Lowe, 2013). This database served as the basis for the data collected for Clayton and Thomson’s (2016) study investigating the CDFs effect on conflict intensity. By analyzing the news media used by Carey, Mitchell and Lowe (2013) to code the PGMD, Clayton and Thomson (2016) identified which of the identified PGMs who could be classified as CDFs. When coding, they found this particular type of militias occurring beyond the scope of the PGMD, and registered CDFs
within the timeframe of 1972 until 2007. Clayton and Thomson’s (2016) dataset is the basis of information for the independent variable in this study. However, the sample used here maintains the timeframe of the PGMD. This, since Clayton and Thomson (2016) merely backtracked the PGMs found in the PGMD. The study found a wider scope than the years 1981-2007. However, it is possible that additional CDFs than the ones found and backtracked were active between 1972 and 1981, but not identified as only qualitative work was conducted in search for classification of the militias included in the PGMD.

In order to have variation in the independent variable, the sample selection includes all intrastate armed conflict episodes (as coded by the UCDP) occurring during the time period of 1981-2007. The length of an episode is determined by the number of recorded battle-related deaths. Limiting the sample to observations taking place the years 1981-2007 with at least 25 battle-related deaths, the sample size of this thesis amounts to 253 observations.

The scope of the theory limits the generalizability to intrastate armed conflicts even though rational choice theory is applicable to both inter-, and intrastate conflict. Information about Pro-Government Militias have been collected in the PGMD with a domestic focus with the aim to analyze and understand why and how governments domestically outsource their monopoly on violence (Carey and Mitchell 2017:128-129). The theory for this thesis builds on the same foundation, narrowing the focus to the effects of a specific type of government militia. Consequently, as empirics and theory does not include international conflicts, the scope of the theory is limited to internal armed conflicts. Furthermore, the definition of Civilian Defense Force emphasizes its local connection. The access to information they are assumed to have, rests on this assumption. The empiric data is collected from the entire population of intrastate armed conflicts without consideration of, for example, geographic location or type of incompatibility. The theory will then be generalizable to all intrastate armed conflicts in a modern world setting, giving good significance to the study.

3.3. Dependent Variable – Duration

The dependent variable is measured as a continuous variable. Duration is operationalized as the number of months in a conflict episode. Measuring in months, instead of years, allows for a more disaggregated analysis of the conflict episodes duration as a conflict episode lasting for one month, or a couple of months, will receive a different value than conflicts lasting a whole year. A conflict episode is defined in accordance with the Uppsala Conflict Data Program (UCDP) as a period of continuous conflict activity. The start date is recorded the date when an episode reached 25 battle-related deaths and continued until the date where the continuation of violence has ceased. A new conflict episode is then recorded if the conflict restarts after one or more years of inactivity.
Thus, the same intrastate conflict can be recorded multiple times as different observations if the conflict activity is non-continuous and reaches 25-battle related deaths. Furthermore, conflicts within the same nation regarding different incompatibilities are recorded as different observations even though they might be active during the same period in time (Gleditsch et al., 2002). The purpose of this separation is to give an as fair measurement as possible of violent conflicts duration. If the conflicts were to be measured from the start date of the first violence and recorded as continuing until the termination of all violence related to the incompatibility, or until conflict resolution, it is possible that low intensity conflicts which remain inactive for an abundance of years are given a high value in duration and therefore bias the results (Cunningham 2006:881). However, conflicts lasting less than a month or more than a month, but not up till two months, are given the same value of 1 month. Giving conflict episodes the same value even though they could differ with up till seven weeks can be problematic. This obstacle is acknowledged as a shortcoming in the operationalization and measurement of the dependent variable. In order to overcome this issue, the variable should be disaggregated further, suggestively in days. However, this is deemed as unnecessary for this study since the aim is to establish correlation in an unexplored field giving future research the possibility to further explore the details of the possible correlation.

3.4. Independent Variable – Presence of CDF

For the purpose of this study, the independent variable is coded as a dichotomous variable with the value of (1) if there existed one or more CDF(s) in the conflict during the conflict episode, or a value of (0) if no CDFs were recorded. The variable is operationalized according to the three criteria defined by Clayton and Thomson (2016) when identifying CDFs in the Pro-Government Militia Database. The three criteria concern the members of the group, their area of operation as well as their role in the conflict, formulated the following way:

1. Participants: the group is composed mainly of civilians.
2. Area of operation: forces remain within the neighborhood, village, or region/department from which they were recruited. Members remain in their own homes rather than reporting to a centralized operational base.
3. Defensive role: civilians mobilize for neighborhood/village security and limited defensive capabilities against insurgents. This can include intelligence gathering and denunciations of insurgents and insurgent sympathizers present or living in the area of operation. It can also include direct combat with insurgents in expelling members of the insurgency from the civilian defense force community or denying insurgents access.  

(Clayton and Thomson 2016:502-503)
All three criteria need to be met in order for the militia to be accounted for as a CDF and included in the Clayton and Thomson dataset. From this, the information was coded into this thesis dataset in the conflict episode it concerned. The conflict episodes which did not have a recorded CDF were coded as a (0). In general, the presence of CDF is coded as a (1) if the conflict episode overlaps temporally with the existence of the CDF and the country only has one active conflict, as it is assumed that the CDF was created in relation to the conflict. If however, a country experiences a number of conflicts over a number of episodes, qualitative work was conducted with the help of PGMD in order to specify the geographical placement of the CDFs (Carey, Mitchell, and Lowe 2013b). This was then matched to the UCDP recorded geographical placement of the conflict. Conflicts occurring in the same area and time period as the CDF meant that CDF presence was registered (coded as 1) to the relevant conflict episode. This qualitative work was, for example, necessary for India where a number of conflict episodes coincided with CDF occurrence. All of the country’s CDFs were recorded by the PGMD to be active in the western provinces of Punjab and Kashmir. Thus, even though some CDFs matched temporally with conflicts in Manipur and Tripura, these conflicts were not coded as having CDF presence since these provinces are located in another part of the country. Another issue related to Myanmar, which proved to be one of the toughest cases to code. The single recorded CDF was documented as active in the “western regions” of the country during an extensive time period (Carey, Mitchell, and Lowe 2013b) when multiple conflicts were active. The geographic placement was deemed too imprecise to code based on geography, and therefore the geographical information was overlooked and coding was made only in relation to the temporal aspect.

Just as Cunningham et al (2009:592) argue in their study, analyzing intrastate civil conflict at the country level could mean a number of aggregations biases. Since this study has collected its information on the existence of CDFs on country level and attempted to disaggregate it to conflict episodes on conflict level, there might be unsystematic errors in the data concerning the geographic placement of the CDF in relation to the conflict. As explained above, this problem has been handled through some qualitative research. However, this research does not cover all possible shortcomings of the insufficient data. Therefore, it is generally assumed that the CDFs included based temporal correlation were active in the area of conflict relevant for the conflict episode.

3.5. Control Variables

The control variables are chosen based on previous research on duration of intrastate armed conflict. All control variables have been found to affect duration of civil armed conflict and have therefore been included to be controlled for in order to, as much as possible, isolate the possible effects CDFs might have on duration.
In their study, Collier and Hoeffler (2004), proxy the effects of greed on internal armed conflict through the presence of natural resources. Collier et al (2004) also recognize economic incentives for conflict, as they find that low per capita income is one of the primary reasons of conflict duration. Therefore, this study includes a control variable of economic inequality as a proxy for greed. As unevenly distributed resources could drive greed, the variable is operationalized through the Gini coefficient’s distribution of income. This data is collected from the VDem dataset where the measurement ranges from (0) to (100) where (0) is perfect equality and (100) is perfect inequality (Investopedia, 2008; Coppedge et al. 2017).

Societal inequality relating to political repression is included as a control variable, as it can be a driver for both greed and grievance in a cost/benefit analysis. Inequality can make individuals or groups fight for greater benefits, either out of grievance from the injustice, or greed for the possible gains (Brown and Langer 2010:29). Societal inequality is a thick concept here conceptualized through the possibility of political representation. Political representation is measured with two categorical variables; power distribution based on economic position and power distribution based on social group. Both variables have values from (0) to (4) where a (4) indicates that the country has an equal distribution of power no matter economic or social group affiliation and a (0) indicates that only elite’s in society hold political power (Coppedge et al. 2017:259-260).

State capacity has been studied as a variable affecting conflicts’ duration. One indicator of a state’s capacity is the strength of the government forces. This has been found to affect both conflict outcome and duration (DeRouen and Sobeck 2004). Therefore, the strength relationship between the rebel and the government forces for each dyad is measured. The dyad accounted for is the primary rebel group for each conflict episode in relation to the government forces. For this control variable, the Non-State Actor (NSA) dataset’s coding is used where the strength is measured as a categorical variable in five ordinal categories ranging from much weaker to much stronger (Cunningham, Gleditsch, and Salehyan 2009:4).

Data on the variables concerning relative strength, social repression and economic inequality are all collected in the year prior to conflict outbreak as this is standard for studies on conflict duration (DeRouen and Sobeck, 2004:308). Through this, the most accurate measure of the situation perceived by the actors is captured.

When a conflict consists of more than two parties, the bargaining range, the sharing of private information and the cost/benefit calculations are affected. Agreeing over an issue is more difficult when more than two parties need to be satisfied (Cunningham 2006:891). Thus, multiple parties may have a major impact on the presented theory as these conflicts should be longer. Hence, the presence of secondary parties is controlled for using the UCDP dataset. The variable is here measured
through coding a continuous variable. The UCDP operationalize a secondary party as an independent state sharing the conflict’s incompatibility, supporting with troops on either side of the conflict (Gleditsch et al., 2002). In order to capture a wider conception of multiparty conflicts this was coded together with additional rebel groups in the conflict episode. A simple dyad is coded with a (1), for every additional party in the conflict, be it a secondary party or another rebel group, one digit is added to the value of the data point.

Different types of third party intervention has been found to affect conflict duration (Regan 2002). Therefore, a dichotomous variable was included measuring whether a UN intervention was present in the country during the relevant conflict episode.

The access to information for both the government and rebels can be greatly influenced by the sheer distances between the central power and the rebels. The size of the country could have real impact on the access to information (Walter 2009:250). The placement of the locus of conflict affects the duration if the locus is far from the center of power (Fearon 2004:287). Therefore, country size is controlled for by logging the size of the country in kilometers squared (divided by 100 for smaller numbers). Data on country size was collected from The World Bank (2017).

To summarize, the table (1) below depicts the distribution of the data on all variables included in the analysis. As the table illustrates, the variables Income Inequality and Reb/Gov Strength are the only variables without data for all the observations in the sample. This will be further discussed and handled in section 4.3. Furthermore, the distribution of the IV and DV will be further discussed in section 4.1.

<table>
<thead>
<tr>
<th>Duration (DV)</th>
<th>CDF Presence (IV)</th>
<th>Income Ineq.</th>
<th>Pol Rep Eco.</th>
<th>Pol Rep Soc.</th>
<th>Reb/Gov Strength</th>
<th>Sec Par</th>
<th>Thi Par</th>
<th>Cou Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>1</td>
<td>0</td>
<td>16.23</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.861</td>
</tr>
<tr>
<td>Q1</td>
<td>1</td>
<td>0</td>
<td>31.80</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>176.520</td>
</tr>
<tr>
<td>Median</td>
<td>16</td>
<td>0</td>
<td>37.55</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>653.640</td>
</tr>
<tr>
<td>Mean</td>
<td>54.02</td>
<td>0.25</td>
<td>39.01</td>
<td>1.992</td>
<td>1.50</td>
<td>1.714</td>
<td>1.348</td>
<td>0.1976</td>
</tr>
<tr>
<td>Q3</td>
<td>59</td>
<td>1</td>
<td>44.00</td>
<td>3.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.00</td>
<td>1259.200</td>
</tr>
<tr>
<td>Max</td>
<td>561</td>
<td>1</td>
<td>65.78</td>
<td>4.00</td>
<td>4.00</td>
<td>5.00</td>
<td>13.00</td>
<td>1.00</td>
</tr>
<tr>
<td>IQR</td>
<td>58</td>
<td>1</td>
<td>12.20</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>253</td>
<td>253</td>
<td>209</td>
<td>253</td>
<td>253</td>
<td>241</td>
<td>253</td>
<td>253</td>
</tr>
</tbody>
</table>

Table (1) The table shows descriptive statistics the dependent, independent and all control variables

3.6. Validity & Reliability

The data for the independent variable is collected from the dataset created by Clayton and Thomson (2016) for their study on CDF effects on conflict intensity. The study collected its data on CDFs using the PGMD as source. From there, the primary news sources used by the PGMD in search for PGMs was used to disaggregate the information. Through this, they were able to
discern which of the militias coded by the PGMD as PGMs were CDFs in accordance with their definition of the phenomenon. In the cases where the online relational database provided by the PGMD did not have sufficient information, the scholars searched for information in major news sources such as BBC (Clayton and Thomson 2016:503). The data from this dataset is transparently accounted for and coded with clear connection between the conceptual definition, operationalization and measurement creating both good validity and reliability. For this study, the same phenomenon (CDF) is used to explain a different outcome variable and therefore the definition is maintained. However, in the transition of the data to this study, both systemic and random errors might have occurred affecting the reliability of the measurements.

Unsystematic errors are errors which deviate randomly from the correct measurement. As discussed above in the measurement of the independent variable, incomplete information required the author to make coding decisions based on temporal overlap rather than complete information on the CDFs geographical presence in some of the conflict episodes. The majority of the observations should theoretically be correctly coded, however unsystematic errors is a possibility when dealing with incomplete data making the coding in relation to reality out of control for the researcher. The most efficient way to overcome unsystematic errors’ effect on the results, is through enlarging the sample size. With a sample size of 253 observations, the most damaging errors are believed to be surpassed.

Furthermore, measuring the independent variable of this study as a dichotomous variable creates validity issues necessary to address. Due to the bluntness of the measure, in some cases the CDF is given more explanatory value than what might be realistic. Therefore, the measurement might not in all aspects capture the concept of “CDF presence” which is attempted to be measured. However, as this is an explorative study analyzing the explanatory value of this relatively newly conceptualized phenomenon’s effect on duration, the bluntness of the measurement is an acceptable shortcoming in validity in relation to the objective.

Systemic error may occur in the measurement of all variables as a consequence of using datasets prepared by others. Systemic error implies that the measurement of all observations deviates in the same direction creating an error which is not easily detected. These issues are issues a researcher need to be aware of when analyzing data and results. However, as long as the raw data is not collected through individual field work, the sources of systemic errors are hard to identify and prevent.

For the dependent variable, little subjective interpretation of its measurement is necessary as it is a simple time measure in months. The background of reporting bias on battle-related deaths is discussed below and regarding reporting bias, unsystematic error is possible. However, beyond
this concern, few problems in terms of unsystematic errors are relevant for this measurement rendering in good reliability of the measure. Furthermore, in relation to the parsimony of the variable, there is little to comment on its validity. Studying duration and conceptualizing it into a measurement of the number of months elapsed during a conflict episode, generates a clear connection between phenomena, operationalization and measurement.

The majority of the control variables intend to capture thick concepts. As only one indicator is used capturing only parts of the concept there are some possible issues of validity in their operationalization. Nevertheless, for the control of political repression as a source for grievance, the validity issue is partially reduced through measuring two different bases for political repression. As for *third party intervention* and *secondary party presence*, data availability limited the measurement to aspects which clearly only capture parts of the concepts generating lower validity. However, the sources of the data used in these measurements are highly reliable fulfilling one of the two criteria on the data.

Much of the data from the VDem dataset is measured as categorical variables, with vague definitions of what determines the value given to each country-year data point. This is also an issue for the categories in the strength relationship measurement gathered from the NSA dataset. Reliability in the measure from reality to dataset is therefore an issue, however for this study the original datasets are openly available, making reliability in terms of replication of this study less acute.

### 3.7. Data & Source Criticism

Data collected from datasets relying on news sources, such as UCDP and PGMD, can always be critiqued. Public news agencies adhere to the public’s interest when choosing what to report about. Consequently, some incidents are reported with vigor while others, deemed as uninteresting to the general public, are overlooked. This means that news close to, or easily relatable for, the western world is more likely to reach English news media, which is the primary source for these datasets. While conflicts or incidents far from the center of interest are disregarded. Furthermore, news media suffer from bias simply when looking at their access to primary sources as they cannot have reporters present in all areas of interest. Therefore, the media relies on secondary or even tertiary sources for reports on events where their own reporters could not be present, creating numerous risk for bias in the information. However, this is a well-known risk when using news media for coding. As these datasets are considered to be the best data available for quantitative studies without individual fieldwork, this author recognizes the inherent risk for bias when analyzing the results, however is forced to disregard it when collecting data.
Missing data has not been a major issue in this study, as most variable information for the individual observations have been accessible in the datasets without the author being forced to extrapolate. The only variable where missing values became an issue, was for the control variable of economic inequality. In the VDem dataset, from which data for the variable economic inequality was collected, the coders extrapolated data using a linear model in places where missing values occurred within a time series (Coppedge et al. 2017:385). Thus, this is out of control for the individual researcher. However, to the extent of this authors knowledge, a linear model removes human subjectivity from the data, once again creating the best situation possible in terms of bias.
4. Results & Analysis

This section will initially present descriptive statistics on the independent and dependent variables. Following this, the results of the regressions will be presented and analyzed. The section will be concluded with a discussion of the main findings and the implications these findings have on the presented theory and research design.

4.1. Descriptive Statistics

Descriptive statistics provide an overview of the variables in the study. This gives indication of the distribution of the data and provides information on potential outliers which can skew the data of the regression giving misleading results.

Shown in Table (2) are descriptive statistics of the dependent and the independent variable of the study. This table shows, amongst other things, the difference in the mean and the median of the IV and DV. The difference between the mean and median of the DV is worth acknowledging. There seems to be a large difference in these two values, indicating extreme values in some observations drawing up the mean but not affecting the median. Moreover, the mean (and the median) is much closer to the minimum value than it is to the maximum indicating a right skew of the data. This can be further discerned in a boxplot (Figure 1). Here, the DV is divided according to its value on the IV; the left box represents the observations with no CDF present and the right box are observations with CDFs present. Figure (1) clearly displays that a

<table>
<thead>
<tr>
<th>Duration (DV)</th>
<th>CDF Presence (IV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>0</td>
</tr>
<tr>
<td>Q1</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>54.02</td>
</tr>
<tr>
<td>Q3</td>
<td>1</td>
</tr>
<tr>
<td>Max</td>
<td>561</td>
</tr>
<tr>
<td>IQR</td>
<td>1</td>
</tr>
<tr>
<td>Sd</td>
<td>93.92</td>
</tr>
<tr>
<td>N</td>
<td>253</td>
</tr>
</tbody>
</table>

Table (2) presenting descriptive statistics on the independent and dependent variables of the study. The values are rounded up to two decimals.

Figure (1) Boxplot of the DV distribution in relation to the IV.
number of outliers in conflict duration are relevant for both conflicts where CDFs have been present and where CDFs have not been present. Furthermore, the mean of the independent variable (Table 2) show that CDF presence is recorded in 25 percent of the sample’s observations. With this in mind, it is easy to see in the boxplot that the observations with no CDF presence are more closely clustered regarding the value of duration, than the conflict episodes where a CDF is recorded. Moreover, Figure (1) shows that both categories, regardless CDF presence, also seem to have a number of outliers. These outliers may be in need of further investigation in order to prevent them from skewing the results.

Other assumptions which can be drawn from Figure (1) relate to the relative distribution of the two CDF categories. The distribution of observations with CDFs recorded have a higher median than observations without CDFs. Remarkably, the median of observations with CDFs is observed to be higher than the value of the third quartile in the opposite category. In general, the distribution of the cases with CDFs (left box) are more dispersed and only two outliers are visually detectable outside of the third quartile. Looking at a histogram (Figure 2) displaying the observations frequency, it can be deducted that a clear majority of the observations of conflict episodes within this sample have a value between 1 and 50 months. This depiction also provides a visual of the possible outliers in the data. Calculating what value of duration which make an observation an extreme outlier\(^1\) gives at hand that all observations with more than a value of 233 months duration are extreme. This excludes 15 observations from the dataset (for a table of these observations, see appendix). These outliers will be included in the primary regressions and excluded for robustness check on the complete model.

\(^1\) Calculated according to the equation \(Q_3+3*IQ\) (TTL, 2017)
4.2. Results

<table>
<thead>
<tr>
<th></th>
<th>Model 1.1</th>
<th>Model 2.1</th>
<th>Model 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF Presence</td>
<td><strong>78.382</strong>*</td>
<td>69.44 ***</td>
<td>65.34***</td>
</tr>
<tr>
<td>Income Inequality (Gini)</td>
<td>0.1765</td>
<td>0.4423</td>
<td></td>
</tr>
<tr>
<td>Political Rep, Eco Group</td>
<td>-11.18</td>
<td>-14.67*</td>
<td></td>
</tr>
<tr>
<td>Political Rep, Soc Group</td>
<td>0.2588</td>
<td>2.169</td>
<td></td>
</tr>
<tr>
<td>Reb/Gov Strength Rel</td>
<td></td>
<td>-13.78</td>
<td></td>
</tr>
<tr>
<td>Secondary Party Presence</td>
<td></td>
<td>-1.490</td>
<td></td>
</tr>
<tr>
<td>Third Party Intervention</td>
<td></td>
<td>31.94*</td>
<td></td>
</tr>
<tr>
<td>Country Size</td>
<td>0.0002574</td>
<td>-0.0001646</td>
<td></td>
</tr>
<tr>
<td>Adjusted R Squared</td>
<td>0.1287</td>
<td>0.1168</td>
<td>0.1251</td>
</tr>
<tr>
<td>N</td>
<td>253</td>
<td>209</td>
<td>198</td>
</tr>
</tbody>
</table>

The p-values are reported according to R studio: p<0.001 is equal to "***", p<0.01 is "**", p<0.05 is "+", p<0.1 is "."

All OLS regressions were made against the dependent variable duration. The models include all outliers.

Table (3) – Depiction of OLS regressions conducted with the full sample of observations.

Table (3) depicts the regression results with all outliers included. Initially a bivariate regression was conducted with the study’s independent and dependent variable (Model 1.1). These results show that there is a correlation between the independent and dependent variable of this study with statistical significance at the 99.99% level. This indicates that the result is generalizable to the entire population of cases, and in 99.99% of an infinite number of times a sample is extracted from the population, similar results will be generated. The relationship is positive indicating that there is an increase in conflict duration of 78 months if the conflict changes from having no CDFs present to CDF presence. As a second step, in order to study the effect the incompatibility and the motivations for conflict onset has on conflict duration, a multivariate regression was conducted when controlling for background factors for conflict onset (Model 2.1). The results for the IV and DV continue to show a statistical significance of 99.99% with a positive effect. The effect size is somewhat diminished; however, it still indicates an increase in duration of 69 months when a conflict changes from no CDF presence to CDF presence when controlling for the background factors. None of the control variables are calculated to have statistical significance. Consequently, these results cannot be generalized to the population. Still to be mentioned, as Model 2.1 show that income inequality and political representation related to social group both estimate small positive effects on duration. However, the results are less than one month. Furthermore, the control variable of country...
size is positive, however the effect indicates a negligible effect of country size on conflict duration when controlling for the other variables. Political representation based on economic group is the only control variable in Model 2.1 with an estimated effect of more than one month. The effect indicates a negative relationship of 11 months for every step on the ordinal scale from complete inequality based on economic group towards complete equality regardless of economic group. Noteworthy for Model 2.1, the number of observations have been greatly diminished from 253 to 209 observations included in the regression. This is due to the control variable derived from the Gini coefficient (Economic Inequality) where missing data for the relevant years became an issue.

Lastly a regression with all control variables of this study was conducted (Model 3.1). This model indicates the estimated effect of CDF presence on conflict duration as 65 months at the same level of statistical significance as the earlier models. Evidently, the estimated effect is once again somewhat diminished, however still substantial. The hypothesis presented in this thesis proposed that CDF presence should decrease the duration of conflict. The effect found in these models estimate the opposite effect. With high values of statistical significance, the implicit null hypothesis can be rejected since there is correlation between the IV and DV, however the results indicate that H1 is incorrect.

Furthermore, in this model (3.1) two additional variables show results with a 95% level of statistical significance which give them explanatory value towards the entire population. Compared to the previous model (2.1), political representation based on economic group is once again estimated to have a negative effect on conflict duration, yet for the full model where controls for factors relating to conflict dynamins is also included, the results show significance. The effect size is also increased. Once again, the sample size is reduced due to missing values in the variables. It is possible that the changes in effect and levels of significance of the different variables is related to the observations removed rather than the added control variables and should therefore be controlled in order to establish their accuracy (see next section).

Third party intervention shows a statistical significance at the 95% level with a strong positive effect on duration (Model 3.1). However discouraging for peacekeeping idealists, this finding is in line with previous research where it has been found that any type of third party intervention causes an increase in the duration of armed conflict (Regan, 2002:71). Still, in this study only interventions by the UN is included in the coding. Generally, UN interventions are intended to be politically neutral (UN, 2017) and findings from Regan’s study excludes international organizations from the findings on neutral interventions when stating that neutral interventions have been found to be strongly associated with longer conflicts. According to the study, interventions by international organizations rather than neutral nations should not independently affect conflict duration (Regan
2002:72). Thus, the evidence from this study shows a different result.

Even though not estimated to have statistical significance and therefore not to be generalizable to the population, political representation by social group experience an effect amplification from less than a month to slightly more than 2 months. Income inequality, and country size are continuously of negligible effect on the dependent variable in the full model (3.1).

Regarding the factors relating to conflict dynamics, both the strength relationship between governmental and rebel forces and secondary party presence estimate a negative effect on conflict duration. In other terms, the stronger the rebel group in relation to the government, the shorter the conflict, and each additional secondary party in the conflict is estimated to decrease the duration of the conflict. The effects of force strength relationship and duration corresponds with previous research on the subject, where it has been found specifically that strong rebels result in short conflicts with a high probability of rebel victory and weak rebels result in longer conflicts with low probabilities of government victory (DeRouen and Sobeck, 2004). However, as previously mentioned, the result of this regression (Model 3.1) does not give statistical significance and can therefore not be claimed to give generalizable support to these findings. Previously, multiple parties in a conflict have been found to deepen the information asymmetries and augment difficulties in finding bargaining space (Cunningham, 2006). Surprisingly, the results in Model 3.1 indicate the opposite, estimating a marginal negative effect on conflict duration for each additional party. Nevertheless, as mentioned previously, the lack of significance render these results ungeneralizable to the general population.

The Adjusted R Squared (Adj R2) is to be mentioned in regard to the models presented in Table (3). The Adj R2 gives a percentage value for the goodness-of-fit for the calculated model. The value punishes the model for every additional variable included in order for the researcher to assess whether the sheer increase in variables has an impact on the result rather than the actual value of the indicators the variables aims at measuring. The values for Adj R2 in Table (3) can be considered constant even though the value of Model 2.1 shows a decrease in the goodness-of-fit with 1 percent. Hence, the chosen variables can be considered to be relevant for the studied correlation.

4.3. Control for missing values

In order to control that the effectual change between the IV and DV is not due to the substantial reduction of observations, but related to the effects of the variables, a regression test was conducted with alterations in the variable Income Inequality. This variable was chosen since the majority of the missing values which create the reduction of the sample size, relate to this variable (see Table 1). Table (4) depicts the regression results when the inequality variable was transformed into a dummy where the value (1) was given to the observations with missing value and the value
(0) was given to the observations with a recorded value. This modulation of the variable gives explanatory value to the previously excluded observations and no explanatory value to the observations with recoded data on the variable without excluding any observations. As the level of significance in the independent variable (CDF Presence) is not altered when all of the missing observations are given a value (and therefore not omitted in the regression), it can be concluded that it is not the omitted observations which have an effect on the result but rather the control variable which cause the effectual difference.

4.4. Robustness Checks

As concluded by the descriptive statistics, there are outliers in the data which can misguide the effects of the results. Therefore, a robustness check of the results was conducted by removing the 15 observations which have been calculated as extreme outliers. In order to properly study the difference in the results, regressions of all models from the original sample were repeated without the extreme outliers. As Table (5) depicts, the results for the estimated effect of CDF presence has diminished with more than 20 months for the full model without impacting the level of statistical significance (Model 3.2). As can be seen when comparing Model 2.1 (Table 3) and Model 2.2 (Table 5), the effectual difference on the DV in relation to the IV is substantial when controlling for background effects with robust variables.
Similarly, there is a stark difference in effects between both the bivariate models (Model 1.1 and 1.2). Regarding the background factors, the effect of both variables relating to societal inequality have diminished notably with the robust model. As the observations removed are the conflicts which have endured for the longest, it is not surprising that factors relating to societal inequality show diminished effects. It is natural to assume that conflicts which were initiated a long time ago are more likely to have low values in societal equalities since these types of factors have developed together with the industrialization. Then, as these observations are removed, the effect is diminished. Another striking difference between the two samples relates to the control variable Third Party Intervention. This control, which is significant at the 90% level in the original full model (Model 3.1) increases in significance to the 99.99% level and the effect size increases with 9 months controlling for the other variables (Model 3.2), indicating that the estimated effect of third party intervention on conflict increases the duration with 39 months. Other changes in the control variables worth highlighting, is the significance gained (albeit at the 90% level) for Rebel/Government Strength Relationship. The effect size on the Rebel/Government Strength Relationship is somewhat reduced, however the results in Model 3.2 estimates a reduction of conflict duration the stronger the rebel group’s forces are in relation to government forces with generalizable results. It should
also be noted that the Adj R2 in Model 3.2 (the full robust model) shows the highest value of the presented models, indicating that the values of this regression might be the best representation of the data.

For further robustness checks, the same regressions were conducted when log-transforming the dependent variable. Simply explained, transforming the dependent variable with this method emphasizes the explanatory value of the many observations with low values on the duration variable and gives less explanatory value to the few observations with high values in duration. In other words, this is a different method for testing the robustness of the data without removing the extreme outliers. The results of this regression can be found in Table (6). At first glance, the results seem vastly different from the full regression models presented previously. However, when using log-transformations the coefficient is to be interpreted differently than in a regular multivariate OLS regression. The results show the estimated percentage increase of a conflict when the variable increases by one step on the relevant scale. For the IV in the full model (Model 3.3) the value then indicates that a conflict increases in duration with 107% percent when the conflict changes from no CDF presence to CDF presence, a very substantial increase which is significant at the 99.99% level. Once again, the full model estimates a substantial increase in duration of third party intervention.

<table>
<thead>
<tr>
<th>Model 1.3</th>
<th>Model 2.3</th>
<th>Model 3.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF Presence</td>
<td>1.2945***</td>
<td>1.17854816***</td>
</tr>
<tr>
<td>Income Inequality (Gini)</td>
<td>-0.01083125</td>
<td>-0.005377063</td>
</tr>
<tr>
<td>Political Rep, Eco Group</td>
<td>-0.16886383</td>
<td>-0.219178658.</td>
</tr>
<tr>
<td>Political Rep, Soc Group</td>
<td>0.09232923</td>
<td>0.102098499</td>
</tr>
<tr>
<td>Rel/Gov Strength Rel.</td>
<td>-0.286897438.</td>
<td></td>
</tr>
<tr>
<td>Secondary Party Presence</td>
<td>0.110226155</td>
<td></td>
</tr>
<tr>
<td>Third Party Intervention</td>
<td>1.242292299***</td>
<td></td>
</tr>
<tr>
<td>Country Size</td>
<td>-0.00001106</td>
<td>0.000002808</td>
</tr>
<tr>
<td>Adjusted R Squared</td>
<td>0.1079</td>
<td>0.07635</td>
</tr>
<tr>
<td>N</td>
<td>253</td>
<td>209</td>
</tr>
</tbody>
</table>

The p-values are reported according to R studio: p< 0.001 is equal to "***", p<0.01 is "**", p<0.05 is "*", p<0.1 is ",."
All OLS regressions were made against the dependent variable duration. The models include all outliers.

Table (6) depicts regression results after log-transforming the dependent variable. All results should be understood as estimated percentage change of the dependent variable.
in the conflict. *Political rep based on economic group* is generalizable, however compared to the original model it loses some explanatory power. Similar to the other robustness check, and contrary to the original full model, *the strength relationship* has significance and shows an estimated a negative effect.

4.5. Discussion

The implicit null hypothesis of this thesis is rejected as all results give statistical significance to a correlation between CDF presence and internal armed conflict duration. Thus, there is a statistically established correlation between CDF presence and conflict duration. The hypothesis of this thesis (H1) does not seem to be supported since the results of all models estimate a positive relationship between the two variables, indicating that conflicts where Civilian Defense Forces are present are longer than conflicts without them. However, the values of Adj R2 in the regressions indicate that the variables studied in this research do not tell the full story. Adj R2 ranges between 0 and 1 and can be understood in how much explanatory value the model has (i.e. how many percent the model explains). Looking at Model 3.2 in Table (5), the Adj R2 has a value of 0.1939, which is highest value of all the tested models. This indicates that this model represents the best explanation of the variance between the variables. Even though it is the highest value in the analysis, the level of Adj R2 indicates that more factors are relevant for explaining the true reasons for why some conflicts endure.

As previously discussed, the causal mechanism, reasoned to lead to shorter conflicts if CDFs are present, relates to the reduction of information asymmetries. Previous findings, which this study builds on, have found that CDFs do in fact provide essential intelligence to governments which substantially increase their possibility of winning (Peic, 2014). Though the findings of this study indicate an increase in duration due to the presence of CDFs, methodological shortcomings might mean that the captured data and the results are not able to test the intended mechanism. CDFs are a subcategory of PGMs and PGMs have previously been found to have different types of effects on conflict continuation and the result of the conflict. It is not uncommon that different types of PGMs are present in the same conflict. This is an occurrence hard to discern as irregular forces do not attract as much attention as regular forces and primary actors (Carey and Mitchell, 2017:134). The research design of this study does not include a measurement which captures other types of militias. It is possible that other militias are more prominent in conflicts where CDFs occur. The effects of other militias might diminish the effect of the CDFs.

Furthermore, there are aspects overlooked in the theory which relate to the local connection of the CDF. Being connected to the state, the members of the CDF can be held accountable for the government’s actions through local social networks. In a society where they share identity and/or ethnicity with insurgents, their relationship with the state might become problematic (Carey
Undercutting the rebels locally can cause the CDF to become the primary enemy for the rebels (Jentzsch, et al., 2015:760). When the militias are targeted instead of the government the conflict can be interpreted as less costly for the government than previously, giving less incentives for the government to invest in the conflict. Keeping this aspect in mind, with the information provided by the CDF, the government may be able to terminate the conflict but choose not to, as further involvement in the conflict might generate higher costs than to let the CDF take the hit.

Another aspect to be considered relates to the local militia’s ability to disturb the conflict dynamics while contributing with information to the state. As the findings of Clayton and Thomson (2016) emphasize, CDFs decrease government use of indiscriminate violence but create an even greater increase in violence from rebels. This is understood to be due to splintering within the rebel groups as well as because of revenge actions. The theory presented in this thesis claimed that this could be a catalyst for the conflict to find bargaining space. However, the effect of the information contribution could be diminished, or altered, due to confusion about the situation amongst all parties generated by local increase in conflict intensity. Then, the mechanism relating to multiparty conflicts (Cunningham, 2006) may be what the results in this thesis indicate, rather than the effects of the CDFs information contribution.

Further, regarding the measurement of the IV, it is possible that the government receive accurate information about the rebels once the CDFs are deployed, however CDFs might be more likely to occur if the conflict has been ongoing for a long time. Then, the government might receive intelligence which contributes to conflict termination, however as the conflict has already endured for a long time, the effect of the CDF is not possible to discern when the research is designed with a dummy IV. Therefore, a recommendation to future research is to measure the occurrence of CDFs more precisely in order to detect possible endogeneity. Furthermore, an interesting aspect of CDF presence not accounted for in this study relates to the dynamics of the conflict if the CDF was present in the area before conflict outbreak. How does the conflict dynamics change if the government has outsourced their monopoly on violence before conflict?

The strength of the militia is also unstudied in this research design. As Jentzsch, Kalyvas, and Schubiger (2015) highlight, a key hypothesis which remain to be studied is how the government’s ability affect the militia relate to the duration of the conflict. It is theorized that strong governments with local militias might impact the conflict duration differently than weak governments who use militias. Here, the emphasis is put on the relative strength of the militia and the government. This is an aspect not taken into consideration for the research design of this thesis where only the relative strength of the government and rebel group is controlled for. Furthermore, if the strength
relationship between the government and the militia can be relevant for duration, the strength relationship between the militia and the rebel group should plausibly be similarly relevant. In some conflicts during periods of time, it is possible that the only present representation of the state within an area is the CDF. With this, the CDF can become the primary target of the rebel group. Therefore, the balance of capabilities between the CDF and the government becomes relevant in order to understand how much influence the government might be able to exert on the CDF and the capability balance between the CDF and the rebel actor is necessary to understand in order to determine how the rebel group might act against the government. Previous findings on strength relationships between the conflicting parties show the same relationship found in this study. Weak rebels are less likely to win conflicts as they are unlikely to gain military victory and the government has little reason to give concessions to a rebel group which is not causing unbearable suffering (DeRouen and Sobeck, 2004). The stronger the rebel force becomes in relation to the governmental forces, the shorter the conflict. If the CDF is the representation of the state and the primary target for the rebels, the CDF can be considered the state actor. The relative strength of the primary actors (CDF and rebel group) may therefore serve as a stronger mechanism than the mechanism relating to information asymmetry between the state and the rebels. Therefore, previous literature which has focused on the balance of capabilities as the explanatory mechanism between the conflicting parties may also be for conflicts where CDFs are present. However, focus should be given to another level of analysis than the previously studied dyad of unitary primary parties.

Based on the empirical findings in this study together with previous research, future research should lay emphasis on, and delve deeper into, understanding the dynamics of non-unitary actors. The findings of this study give good reason to further develop scholarly understanding of the actor dynamics at lower levels of analysis. As evidence here shows, the effect of a specific type of sub-actor to the government, can make a difference in years for the duration of the conflict.
5. Summary & Conclusion

The guiding research question of this thesis was aimed at understanding why some intrastate armed conflicts last longer than others. Building on previous findings on duration, a research gap was identified. Earlier, few scholars have focused on analyzing how different features of the government actor might impact duration. The idea of disaggregating the government actor lead the author of this thesis towards the field of Pro Government Militias. There, it is studied how governments’ outsourcing of violence through militias, which are formally or semi-officially connected to the state, impact society. Combining findings in these two fields resulted in a study of a specific type of militia’s impact on duration.

The aim of this paper therefore became to establish whether there is a relationship between the presence of Civilian Defense Forces and the duration of intrastate armed conflict. The proposed theory offers a hypothesis and causal mechanism claiming that CDFs do indeed affect duration and that their presence should decrease the duration of armed conflict. The reasoning conducted is founded in rational choice theory where CDFs are claimed to alleviate the information asymmetries within the conflict, revealing information to the government on the capabilities and motivations of the rebels. This should reveal bargaining space otherwise only found through continued fighting, a process assumed to be costlier in terms of time than information acquisition through CDFs.

Using a sample of 253 observations of conflict episodes between 1981-2007, the study generated results at a significance level of 99.99%, ascertaining that there is correlation between CDF presence and conflict duration. The results found in this study indicate that the presence of CDFs in intrastate armed conflicts have a substantial positive effect on conflict duration.

The findings indicate that the constellation of the government actor has an extensive impact on the duration of conflict. The government actor’s arrangement of forces might be a key dynamic substantially affecting conflict duration. This leads to potential policy implications calling for caution on how and when government actors can, and should, outsource its monopoly on violence. The state is entrusted with heavy responsibilities of protecting its citizenry, in situations of armed conflict, these responsibilities grow even heavier and more complicated to carry. While it has previously been established that CDFs generate more violent conflicts (Clayton and Thomson, 2016), this study further contributes to the field with findings suggesting that CDFs also generate longer conflicts. Together, in relation to the moral aspects of the responsibility to protect, policy makers should thoroughly consider the potentially dire consequences of deploying Civilian Defense Forces in civil armed conflicts.

***
6. References


Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Frida Andersson, Michael Bernhard, M. Steven Fish, and Adam Glynn. 2017. “V-Dem Codebook V7.”


## Appendix

### Table of the observations removed as extreme outliers

<table>
<thead>
<tr>
<th>Conflict</th>
<th>Duration</th>
<th>Conflict Continues beyond the Scope</th>
<th>CDF Presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia vs. KR</td>
<td>238</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Philippines vs. MILF, MNLF</td>
<td>240</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Angola Vs. UNITA</td>
<td>241</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>UK vs. PIKA</td>
<td>243</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>South Africa vs. SWAPO</td>
<td>260</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Turkey vs. PKK</td>
<td>280</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sudan vs. SRF</td>
<td>295</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Philippines vs. CPP</td>
<td>314</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ethiopia vs. EPLF</td>
<td>326</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Afghanistan vs. Taleban</td>
<td>356</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Myanmar vs. KIO</td>
<td>382</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Myanmar vs. CPB</td>
<td>490</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Colombia vs. ELN</td>
<td>516</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Myanmar vs. KNU</td>
<td>526</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Israel vs. PNA</td>
<td>561</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>