Master's thesis

Jo Hallvard Nasvik

Political Regimes and Internal Conflict

NTNU
Norwegian University of
Science and Technology
Faculty of Social Sciences
and Technology Management
Department of Sociology and Political Science

Jo Hallvard Nasvik

Political Regimes and Internal Conflict

A Statistical Analysis Using Automatically Coded Event Data

Master's thesis in Political Science

Trondheim, spring 2014

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU – Trondheim
Norwegian University of
Science and Technology

Jo Hallvard Nasvik
johallva@stud.ntnu.no

Political Regimes and Internal Conflict

A statistical analysis using automatically coded event data

Master's Thesis in Political Science
Trondheim, May 2014
Norwegian University of Science and Technology
Faculty of Social Sciences and Technology Management
Institute of Sociology and Political Science

Abstract.

Democratization has been one of the main policy goals towards reducing internal conflicts in the last decade. However, recent conflict research, and recent events such as the many riots in European democracies, show that the relationship between democracy and internal peace is not as clear-cut as previously thought. This thesis uses new automatically coded event data to perform a statistical analysis of the relationship between regime and six types of internal conflict. I find strong support that democracies see more non-violent conflicts than other regimes. I also find strong support that authoritarian regimes see fewer of any conflicts than other regimes. Finally, I find mixed evidence on democracies and violent conflict; while democracies have a smaller chance of violent conflict initiation, they see more violent conflict events of some types than other regimes.

Acknowledgements.

I would like to thank my supervisor Halvard Buhaug both for posing the question behind this paper and for all his encouragement, help and feedback during the writing process. Without his willingness to review my work periodically and set partial deadlines, this thesis would probably not have been done in time. I would also like to thank Eirik for helping me with the computer technical issues/black magic.

Any remaining errors are solely my responsibility.

Trondheim May 2014

Jo Hallvard Nasvik

Contents

Tables and Figures

1.0 Introduction

In modern times, democracy is often considered a particularly peaceful form of government where internal conflicts are solved through non-violent means. After the Second World War a series of researchers of international relations reassessed the old idea of a republican peace and applied it to modern democracies (Babst 1964, Doyle 1986). This thesis was, in more recent research, extended to internal conflicts, and for a time a partial consensus formed in around the idea of a civil democratic peace (Hegre et al 2001). The relationship of regime type was found to be an inverted- u where both democracies and authoritarian regimes had fewer internal conflicts than anocracies. However, findings in research that has disaggregated conflict, spatial and regime variables have found that this relationship does not hold (Buhaug 2006, Collier & Rhoner 2008, Sobek &Payne 2010). In addition, some major problems have been discovered in the data used to find the inverted-u relationship (Vreeland 2008). One of the key developmental policies of the early 21st century has been the promotion of democratic institutions, but is democracy necessarily as internally peaceful as once believed? Or is perhaps the democratic peace liked only to certain kinds of conflict?

This thesis seeks to analyze the relationship of regime types and conflict types by using new automatically coded event data. The thesis seeks to uncover more knowledge of the relationship between democracy and conflict through disaggregation of conflict types. The research question of the thesis is: *Are some conflict types more prominent in democracies than other regimes?*

The motivation of the paper is twofold. First, it is interesting to uncover more knowledge of the relationship between regime and conflict in light of the wave of protests, riots and other small-scale violence in Europe in the early 21st century (Taylor 2013, Wagner 2013). In addition, the unexpected eruption of The Arabian Spring reiterates the question of why, when and where

political conflicts turn violent.

Secondly, the paper is theoretical interesting. The research front in conflict studies involves a process of bring the state back in as an actor in conflict analysis, disaggregating variables to more precise typologies, and of marrying the civil war and international relations centered conflict research tradition with research on non- violent political conflict. In recent years conflict research has been characterized by a trend of disaggregation of data (Gleditch et al 2014). Conflicts, regimes, state capacity, ethnicity and other important variables in the field have been broken down into typologies that are more precise in order to create a more accurate understanding of casual mechanisms. This thesis explores the possibility of using new automatically coded data to aid this effort. Building on efforts to uncover the relationship between regime types and conflict this paper seeks to disaggregate conflict types in order to see if the same relationship between regime and conflict can be found. Research has revealed that the inverted-u curve of regime and conflict does not hold when civil war is disaggregated or in regard to other types of political violence (Buhaug 2006, Collier & Rohner 2008). In fact democracy has been found to have a positive effect on certain types of conflict, will a similar relationship be revealed if non-violent as well as asymmetric conflict types are also analyzed? This thesis seeks to aid this effort by disaggregating conflict variables and analysis both violent and non-violent conflicts. Even if this thesis finds no interesting relationships between regime and conflict, it should at least discover something about the possibilities for using automatically coded event data to aid variable disaggregation.

In order to answer the research question I construct a data set of conflict event data where a large typology of conflict is aggregated to six conflict types, and tests the relationship between these conflict types and regime using zero inflated negative binomial regression.

The thesis finds strong support that democracies have more incidents of non-violent protests, and that authoritarian regimes have fewer conflicts of any type. Democracy is found to have a mixed effect on some types of violent conflict. On the one hand, democracies have a greater chance than anocracies of avoiding conflict initiation; on the other hand, they experience more events when conflict occurs.

The paper consists of three sections. The first section presents the theory behind the thesis. This section has there objectives; (1) It seeks to review the previous research and the findings from this in order to (2) explain the theoretical part of the motivation for the thesis, and finally (3) to present the theoretical logic behind the analysis. The section begins by explaining the research done on conflict then moves on to consider theories of political conflict. In the end of this section I use the theory and empirical findings presented to craft hypothesis about the expected relationship between regime and conflict.

The second section deals with the method of the thesis and methodological challenges of it. The section starts with describing the statistical method and explaining the reasons why this approach was chosen. It then moves on to consider the data on the dependent variables and show how this data was constructed. Then it will consider the data on regime and on the control variables, before it finally considers the regression models of the analysis and presents descriptive statistics.

The finally section of the thesis contains the findings from the analysis and my discussion of them. The analysis consists of three steps. In step, one the relationship between regime and conflict is considered in a series of binary models. In step two I seek to improve upon this model by including control variables. In step three, I seek to check whether the unexpected findings in step two can be due to methodological problems and to test for curvilinear relationships, interaction and extreme case effects.

Finally, I conclude what my findings are and how this thesis can be improved upon by further research.

2.0 Theory & Research

2.1 Introduction

In this chapter I will review the literature that is most relevant for this paper. The purpose of this is to set the paper in context of this research, in order to make clear its theoretical motivation. There are mainly three sources of research that is examined here: research on regimes, on social movements and dissidence, as well as research on civil wars and other violent conflicts. The reason the I review research and theory form so many different research traditions is that the subject of this paper lies at a crossroads between these research traditions. However, the principal literature that this paper is based on is quantitative research done on governance and conflict after the end of the cold war, particularly research done on civil war.

As this paper seeks to investigate variations in onset of intrastate political conflict this chapter reviews mostly research done on that, and not on duration, severity or interstate conflict.

Before the research review, it is necessary to define some terms used in this paper.
First, regime is central to this paper and I frequently refers to the standard typology of democracy, anocracy and authoritarian regimes. This typology is an oversimplification of the far more complex political reality, however it is sensible for several practical reasons. Since this paper relies on the Polity IV projects data on regime it is only natural that it apply its definitions of the various regime types as well. Polity IV defines democracy as:

> "Democracy is conceived as three essential, interdependent elements.
> One is the presence of institutions and procedures through which citizens can express effective preferences about alternative policies and leaders. Second is the existence of institutionalized constraints on the exercise of power by the executive. Third is the guarantee of civil liberties to all citizens in their daily lives and in acts of political participation (Marshall, & Jaggers 2013: 14)"

Authoritarian regimes are countries that are, by and large, defined by what democratic features

they don't have. They are often referred to as autocracies in the literature, a term that I have chosen to avoid as it can also refer to the rule of one-person specifically. Polity IV defines authoritarian regimes as:

> "Autocracies [Authoritarian regimes] sharply restrict or suppress competitive political participation. Their chief executives are chosen in a regularized process of selection within the political elite, and once in office they exercise power with few institutional constraints. (Marshall, & Jaggers 2013: 15)"

Anocracy is an even more elusive concept that really exists only to brand countries that don't fit the bill of either democratic or authoritarian. They are also often referred to as mixed regimes and frequently thought of as inherently unstable. Polity IV defines anocracy as:

> "Anocracies are a middling category rather than a distinct form of governance. They are countries whose governments are neither fully democratic nor fully autocratic but, rather, combine an, often, incoherent mix of democratic and autocratic traits and practices (Cole & Marshall 2011: 9)"

This paper seeks to investigate the effect of regime on many forms of conflict, however I frequently refer simply to conflict. By conflict I refer to any political conflict, similarly the term dissidence is used to refer to participation in any kind of conflict. I attempt to review research done on all the conflict categories that I use in my analysis, but the majority of research reviewed here comes from civil war research. The reason for this is that civil war research is vary expansive and has lately begun comparing civil war to other conflict types, and disaggregating civil war into more accurate and narrow categories.

This chapter will begin by reviewing what we know of internal conflict from empirical studies. Then review the two major theories that attempt to explain internal conflict, and finally based on theory and empiric findings construct hypothesis about what I will find in the analysis. What conflicts are in the data and how they get there is covered in the methodology section, for now a theoretical definition of conflict will do. The Norwegian Lexicon of Political Science defines conflict as:

"The competition between political actors over goods or values [my own translation] (Østerud 2007: 135)"

This is a quite broad definition, but also quite apt as this paper seeks to explore broad range of conflicts.

2.2 Recent Research and Empirical Research

Instead of dealing with this section chronologically I will start with the article that poses the question this paper is mainly interested in, then I will deal with the literature that forms the basis of it, then the literature that has expanded on it. The reason for this is that the work done in the last 20- ish years is the most relevant to this paper, and the one that best explain the theoretical relevancy and motivation for the paper.

Hegre et al (2001) conduct a statistical analysis of civil war and regime where they conclude that democracies have fewer civil wars than other regimes. The article builds on the previous work done on democracy and conflict that focused on interstate wars, and applied this to civil war. The article found that democracies had fewer civil wars while autocracies had more than democracies but less than anocracies (Hegre et al 2001). The article is based on data from 1816 to 1992, and tests for the effect of regime and regime change on the risk of civil war. They find that civil wars are far more common in anocracies than in democracies and authoritarian regimes. They describe this relationship as an inverted-u. Hegre et al (2001) are not the first to show this relationship, but they are the first to show it in data of this magnitude.

Below I will show first that a similar relationship had previously been found in similar inquiries, and how research in the last ten years has begun to sow doubts of it. The multiple findings of the inverted-u relationship between governance and conflict was taken by many as major evidence of what is known as the democratic peace thesis. This thesis has a long history, some of which will be looked more closely at bellow. It states that democracy as a form of government is based on the idea of providing ways to resolve conflicts peacefully. Therefore, if all governments were

democracies there would be perpetual peace. Immanuel Kant (1795) was the first to develop this idea. He claims that republics will not go to war with each other, as republics would have more to gain from peace and trade (Kant 1795). The extension of Kants (1795) idea of a republican peace to modern democracies has seen some theoretical critique, see Gates, Moses & Knudsen (1996) and Danilove & Clare (2007).

Later Babst (1964) Doyle (1986) and Rummel (1995, 1997) make extensive inquiries into the relationship between regimes and conflict, and conclude that democracies were more peaceful than other regimes. These enquirers were mostly concerned with, and most conclusive, in the case of interstate war. Though their research Kants (1795) idea is applied to modern democracies Babst (1964) made an insightful inquiry into the relationship between regime and interstate war in the first and second world wars. He compared the relative amount of countries that had elected versus non-elected governments amongst the entente (allies) and axis respectively. He found that the allies had overwhelmingly elected governments while the axis had overwhelmingly non-elected governments.

Doyle (1986) preforms a more theoretical work where he reexamines the theoretical works of Machiavelli, Kant and Schumpeter in view of recent research in modern political science. Doyle (1986) thus takes the reformation and enlightenment idea of the republican peace and puts it in context of modern democracy after the Second World War.

Rummel (1995) does a large study of the relationship between regime and conflict were he finds that democracies are over all more peaceful than other regimes. Rummel also (1995: 85) does an analysis the results of as many studies he could find on the subject from before 1980 and concludes that democracies are more internally peaceful.

Maoz and Russet (1993) made an investigation of the relationship between democracy and conflict on the international scene where they find that democracies are about as likely to be involved in wars as other regimes, but are less likely to be at war with each other. They then go on to examine why this could be. They present two different models to explain the negative effect on democracy, and test for both of them. The first model is the normative; which states that

democracies are less prove to conflict between each other because norms of compromise and cooperation are prevalent in democratic societies. The structural model claims that the institutional structure of democracies but constraints on leaders that make conflicts harder to escalate (Maoz & Russet 1993: 1). They find that there is support for both models, but that the support for the normative model is more robust.

Krain and Myers (1997) are among the first to investigate the democratic peace theory in themes of civil war. Inspired by Rummels (1995) work on state violence and regime they hypothesize that democracies should experience fewer years of civil war than other regimes. They conduct a chi-square analysis of democracy and civil war using the correlates of war data set and polity III data on regime. They find that democracies do have fewer civil wars (Krain and Myers 1997: 113).

It is worth mentioning that similar inquiries were made at his time by researchers mainly interested in democracy as a phenomenon rather than political conflict. Powell (1982) makes a large study of all the consolidated democracies at that time, where he investigates into many aspects of them including internal conflict. Many of these results are quite interesting, while they do naturally not illustrate the difference between conflicts in democracies and other regimes, as it makes no such comparison. Powell (1982) finds that most violence in democracies comes from organized groups and is directed at the government or at other political opponents, also he finds that separatist violence can be particularly problematic in democracies and that population size is the largest contributor to violence on its own (Powell 1982:126, 154, 159). Krain (1998) does an updated version of Powells (1982) analysis, where he uses updated analytic methods concerning event counts (that we will return to in the methodology chapter). Krain (1998: 161) finds quite different results using Powells (1982) data; he mainly finds that material and representation problems have a larger impact than grievance problems.

Similarly, some research has been done on variation in use of repression that touches on some of the same dynamics as dissidence, as the two naturally correlate. Findings from this research mirrors findings from research of political conflict, as there has generally been found an inverted – U relationship of regime and repression (Markus & Nesvold 1972, Hibbs 1973, Muller &

Weede 1990, Moore 1998, Eck & Hultman 2007, Henderson & Ragan 2013). As a general tendency both democracies and autocracies repress less than anocracies. This is due to the fact that repression is sometimes met with increased or intensified dissent, and sometimes by decreased dissent. The logic of winch of these two happen is the same as the logic of whether or not dissent should be attempted in the first place. In democracies it is often more useful to engage through legal channels while in autocracies the repression can be so harsh that the fear of it alone is enough to silence dissidents. Anocracies on the other hand fall in between these two extremes in regime type, and therefore repression leads to escalation in violence, which means that by extension the form of political conflict is influenced by regime (Markus & Nesvold 1972, Hibbs 1973, Muller & Weede 1990, Moore 1998, Eck & Hultman 2007, Henderson & Ragan 2013).

It is interesting that such a similar relationship has also been found by those who are looking at the situation from the other side, so to speak. As exceptions to the inverted – U relationship have been found with regard to types of dissidence, could there be similar exceptions in terms of type of repression? That is a question for a different inquiry, as shall be seen later Gurr (1970) says that the natural reaction to violence is violence, and that if this reaction is not alleviated a cycle of violence can quickly form which it is very difficult to break free. In relation to the inverted-u model it seems that consolidated democracies and autocracies have the means to break it while anocracies do not.

Built on this research a partial consensus was begging to form around the inverted-u curve relationship of regime and internal conflict, a civil democratic peace. However, in later research evidence has been found that casts serious doubt on these propositions.

2.3 Economic Development & The Liberal Peace Proposition

Along with democracy, economic development has been the number one policy implication of civil war research after the end of the cold war (Dixon 2009, Enia et al 2011). Gratzke (2007)

takes this argument further and offers up an alternative explanation to the democratic peace where he points to the fact that many democracies are in fact rather wealthy, and presents empirical evidence that the democratic peace is better explained by capitalism. Democracy and peace are both seen as results of economic development. This has in turn been developed into the proposition of a larger liberal peace rather than a democratic one. In this model liberal norms of free market and free trade along with a more narrow definition of liberal democracy rather than simply institutional democracy, are seen to promote peace when operating together (Mousseau 2009, Schnider 2014).

Collier and Rohner (2008) investigate the relationship between wealth, democracy and peace. They find that the effect of democracy is as expected from democratic peace theory when democracy coincides with economic development. However, the effect of democracy was found to be positive on civil war in poor countries (Collier & Rohner 2008: 538). One the other hand economic development is found to make authoritarian regimes more prone to political violence. Furthermore, of great interest to this paper, Collier and Rohners (2008) investigation goes into the relationship between democracy, economic development and *several types* of conflict. They look at riots, coups d'etat, assassinations, demonstrations and strikes. In the case of: riots, strikes, demonstrations and assassinations democracy is found to have a positive effect, for other types of conflict the effect is statistically insignificant (Collier & Rohner 2008: 537). Except when in conjunction with economic development where it has negative effect on all but coups d'états, which effect is statistically insignificant (Collier & Rohner 2008: 537). This is very interesting; not only do they find that democracy has a positive effect on many kinds of *violent* conflict, they also find no negative effect on civil war from democracy. The positive effect of democracy on demonstrations and strikes is to be expected, these things are part of everyday politics in democracies, but the effect on riots and assassinations is unexpected. Democracy is, as shall be seen in the theory section, is often thought of as a means to peacefully resolve conflicts. It is therefore unexpected to find that democracies have more of these kinds of conflicts, and it will be interesting to see if similar results will be found using event data.

Schneider (2014) and Schneider & Gleditsch (2014) review the literature on the capitalist peace. They claim that the discussion is not yet over as there are good arguments on both sides, but that

the capitalist peace has seen substantial empirical support, but that it faces a difficulty in lacking micro level causal links and could benefit from the arrival of 'big data'.
Whatever the link between economic development, capitalism, democracy and peace is, it seems reasonable to me that at this point to assume that there is some link, and it seems necessary to account for the effect of economic development.

2.3.1 Lootable resources

One factor that is often seen in relation to economic development is the presence of lootable resources that can 'fuel' a civil war (Dixon 2009). This is only supported by evidence concerning diamonds and oil (Dixon 2009), and in some research the evidence is less supportive all round. Buhaug & Rød (2006) find that that diamonds have a negative effect on territorial conflicts and a positive effect on governmental conflict. Sobek & Payne (2010) find that lootable resources have a positive effect on civil war onset when they coincide with low government capacity, and have the opposite effect when government capacity is high. This makes intuitive sense, as a stable and capable government might use such a recourse as a means to strengthen themselves, unstable and infective governments would be unable to do so. The resources then becomes an incentive for rebels, who are motivated by greed. The reason for this is that rebels must try to make rebellion more viable for potential recruits by changing the cost/benefit relationship; one way to do this is to offer an economic incentive, often through the use of loot. There is however consensus that oil and civil war correlate (Dixon 2009).

The idea of a liberal peace undermined the argument of the democratic peace thesis as it moves the causality. Institutional democracy is not seen as a cause of peace, but peace and democracy are both seen as results of liberal norms and free markets.

2.4 Disaggregation of Variables.

Several articles published before and after Hegre et al (2001) recognize that civil wars are not a uniform set of events, but wary in their objectives (Collier & Hoffler 1998, Sobek and Payne

2010: 216). Consequently, there has been an effort to disaggregated civil war in later research. This effort has also changed opinions on the democratic peace thesis as it has revealed correlations that contradict it.

As mentioned in the introduction Buhaug (2006) does an analysis of civil war where he distinguishes between civil wars aimed at establishing a new state in part of the current state (territorial), and those aimed at taking over the power in the current country (governmental). He argues that weak states will have more governmental conflicts while strong states will have more territorial conflicts (Buhaug 2006). The basis for this argument is that rebels are rational actors and will adjust their goals to their capabilities (Buhaug 2006: 692). Buhaug (2006) performs an analysis of civil war from 1946 to 1999 measuring the effect of a set of independent variables including regime. The results of the analysis are interesting as they break with the conventional wisdom that democracies have fewer civil war than other regimes. Buhaug (2006: 705) shoves that democracies in fact have more territorial civil wars than other regimes, the effect is also strongest in consolidated democracies. As mentioned in the introduction it is this finding along with the evidence from Hegre et al (2001) that is the basis for this paper; the purpose being to see if evidence can be found, using new event data, that democracy has a similarly different effect on other types of political conflict. Buhaug (2006) hypothesizes and concludes that it is the relative capabilities of rebels vs states that dictate their goals. In relation to regime he finds that democracies have more territorial civil wars than authoritarian regimes, He offers the explanation that while it is difficultly to legitimize a civil war against the democratic regime it shelf, it is possible for democratic regimes to host people of a different ethical or religious identity, who will desire independence (Buhaug 2006).

In the study of terrorism, a similar focus on rational actors has evolved (Sandler 2014). James Piazza (2008) conducts an empirical test of the effect of democracy on terrorism, using country year data created from totaling up the number of events for a country in a year. He finds that democracy and free markets are not significant predictors of terrorism. His study is interesting not just, because it analyses the same relationship as this paper, but it also uses the same methodological approach. However, Piazza uses the problematic Polity and Freedom House indicators that are endogenous to regime, see the methodology section for the details

(Vreeland 2008). Yet he still does not find regime to be a significant predictor of terrorist events. It may be that the problematic indicators are therefore mostly problematic in relation to civil war.

Sandler (2014) writes a review article about the most prominent findings in the study of terrorism. In relation to regime, Sandler (2014) emphasizes that democracies are put in a dilemma by terrorism. If the government fail to react to terrorist threats, they will lose legitimacy because they are seen as incapable of protecting their people. If they do react, it is difficult to combat terrorism without trampling people's democratic rights (Sandler 2014). The US reaction to the attacks of September 11th serves as a good example. The recently revealed National Security Agency's massive surveillance program and the extraordinary powers given to law enforcement through the patriot act are clearly at odds with principles of privacy and being innocent until the opposite is proven. Yet the rationale behind these programs is easy to understand, as without surveillance terrorist plots can be very difficult to uncover, and if they are uncovered they can be difficult to prove. This may cause terrorist to target democracies, as they know they are likely to have effect, as the balance of appropriate response to terrorism can be difficult to find. Similarly democratic freedoms can make the organization and execution terrorist activities easier as it involves freedom of association, communication etc. A similar point is made by Li (2005) who claims that democracy has a positive effect on terrorism in the form of increased freedoms, along with a negative effect in the form of alleviating grievances that cause terrorism in the first place. It may be then that democracies because of this dual relationship could be associated with international terrorism, but not with domestic terrorism.
Building on Buhaugs (2006) analysis Sobek & Payne (2010) expand the typology of civil wars by separating governmental wars into two further categories; (1) wars where the objective was to remove the government (replacement), and (2) wars that seek to change the entire relationship between government and society (legitimacy). Interestingly Sobek & Payne (2010:236) find no statistically significant relationship between regime type and either form of civil war. They do however find significant effects from government capacity, and criticize previous civil war research for representing capacity with a regime variable (Sobek & Payne 2010).

Another aspect of this disaggregation of conflict has been done in recent conflict research is that there has been increasing interest in non-violent forms of conflict, at the same time as

sociological studies of these phenomena are becoming more quantitative (Ackerman & Karatnycky 2005, Chenoweth & Stepahn 2008). Protests, strikes, boycotts and similar non-violent forms of dissidence have been given more attention in the wake of The Arab Spring, The Occupy Movement and the waves of riots in Europe in the last decade. These events have on the on side inspired an interest in the potentially unstable aspects of democracies, and on the other an interest in the nature of non-violent resistance in non-democracies.

Cunningham (2013) analyses violent and non-violent strategies in self-determination conflicts in an attempt to unlock the logic of why some dissidents choose violence and some non-violence. She begins with criticizing that violent and non-violent strategies have been studied as separate phenomenon, and the heavy focus on civil war. She argues that dissidents turn to irregular tactics when they do not achieve their goals thorough normal channels, or if no normal channels exist (Cunningham 2013: 294). Groups choose violent tactics when they seem more effective than non-violence. Additionally she points out that violent tactics, especially terrorism, requires fewer people than non-violent tactics, and is therefore chosen when dissidents are few in number. She finds that demand for independence, political exclusion and economic discrimination have effect on both violent and non-violent strategies. While democracy and state capacity measured in GDPPC, relative group size, group concentration and kin in adjoining state have different effects on civil war and non-violent champagnes (Cunningham 2013:301). On the basis of this she argues that civil war and non-violent strategies should not be considered as entirely separate phenomenon, as they are affected by some of the same variables (Cunningham 2013: 302).

2.4.1 Separating Regime from Regime Stability and State Capacity.

Another area where conflict research has moved towards increased disaggregation is operationalization of regime. Particularly there has been an effort towards separating the regime type from regime consolidation, stability and capacity. As Hegre et al (2001) note this is an important distinction, as anocracies may be more prone to conflict also because they on average have gone through transitions more recently.

Cederman et al (2010a) are critical of regime being operationalized as a static variable, and employ an algorithm that can locate periods of regime change in the polity data to assess the link between democratization in addition to regime on conflict. They find that democratization, and to a lesser degree autocratization, is associated with civil war. This relationship is also distinct from the curvilinear relationship between current regime and civil war, but is only present in the case of governmental conflicts (Cederman et al 2010a). Later Cederman et al (2013) add to this by examining the relationship of competitive vs non-competitive elections and civil war. Their argument is that elections should be associated with conflict in unconsolidated democracies because elites have to mobilize voters and this is most easily done on the basis of preexisting lines of division and conflict in society. Also there is a risk that the looser of an election will not except the defeat and seek violent means to power after the election (Cederman et al (2013). The find that elections are associated with conflict, mainly in relation to ethnic violence.

However previous research on the relationship between regime instability and civil war by Hegre et al (2009:188) shows that the inverted u-curve of regime and civil war holds even when testing for regime instability. Though the results in this analysis are not entirely robust concerning onset, and much of the effect of democracy seems to be somewhat linked to wealth and stable institutions.

In a previous investigation of the effect of institutional inconsistency and political instability Strand et al (2006) using Gurr's (1970) polity data set to examine changes in institutions. They found that institutional inconsistency had a large effect on political stability, and claim that semi-democracies do not only experience more civil war as a result of being prone to invoke grievances, but also because of weak institutional capacity to respond to dissent (Strand et al 2006: 907).

Gleditsch and Ruggeri (2010) argue that using regime variables as proxies for political opportunity is problematic, and attempt to create different measurements of opportunity. They use data on leader entry as a measure of political opportunity. Their argument is that irregular

leader entry is a sign of institutional inconsistency and a good proxy for political opportunity. They show that there are (in their data) far more episodes of irregular leader entry than there are episodes of change in polity score. This shows that polity works poorly as an indicator of institutional instability because a country may have unstable institutions without changing regime characteristics (Gleditsch & Ruggeri 2010). They find that irregular leader entry is associated with civil war, they also find that the effect of democracy on civil war is negative when political opportunity is accounted for (Gleditsch & Ruggeri 2010).

Theis (2010: 325) adopts a more economical measurement of state capacity where he calculates complex indicators of the size of the government in the economy as well as the relative capacity of that government, compared to governments with similar economic development and resource endowment, in terms of tax collection. He finds that state capacity has no effect on civil war onset, and neither does primer y product export (Theis 2010: 321, 327-8).

In view of these articles, it seems clear that if one is to isolate the effect of regime on political conflict one must account for the effect of institutional stability and governmental capacity. However, in what manner state capacity is to be operationalized is a question that still need answering in the literature.

2.4.2 Disaggregating The Spatial Dimension.

A final area where there has been considerable efforts to disaggregate variables is in the spatial dimension. The importance of geography in explaining civil war has been acknowledged for some time. Several articles of importance have used measures of difficult terrain as an indicator for civil war and especially guerrilla tactics opportunity (Collier & Hoeffler 2004, Fearon & Laitin 2003). However, the spatial scope of most conflict studies have either used country year data (the monadic level) or focused on a single case. This is not the best way to analyse internal conflicts as many of them involve only a small part of a given country (Buhaug & Gates 2002). Disaggregating the spatial dimention of conflict therefore means to analyse conflict on a sub-

national level, and providing data on ethnic groups, economic realities and other variables know to explain conflict at that level (Buhaug & Gates 2002, Buhaug & Lujala 2005, Buhaug 2010).

2.5 Other Notable Findings

A large number of variables have been found to correlate with one conflict type or other. There is no need to review them all here, as not all of them are significant for many types of conflict, nor is there space for it. However, there are some that need to be mentioned.

Ethnicity was long thought to be one of the top, if not the, explaining variable of civil war. With the breaking up of old European empires into countries whose boarders were either drawn by ruler, or even worse, deliberately set to dived the population in order to keep them from uniting against the colonial power. Moreover, with the collapse of The Soviet Union in the early 90s, ethnicity became a central part of many intrastate conflicts in the late 20th and early 21th century (Gurr 2000). The body of work on ethnicity is massive and there is only space to review the most relevant articles here. As for other types of diversity, such as religious diversity Dixon (2009: 710) notes societies are split along many such lines, but little consensus exists on how they ought to be operationalized.

The Minorities at risk project (MAR) is probably the most extensive work done on ethnicity and political conflict. It is a vast project that collects information on, and analyzes ethnic minorities who are politically active (Minorities at Risk 2014). As well as a number of publications, the main product of the project is the MAR dataset, which covers 283 groups form 1945 – present (Minorities at Risk 2014).
Ellingsen (2000) reviews previous work on ethnicity as well as other variables that affects civil war, and conducts an analysis of it. Ellingsen (2000:244) finds that multiethnicity is associated with conflict, but does not find support that the size of the minority has impact on conflict onset. Fearon and Laitin (2003) later conduct one of the most famous investigations into the relationship of ethnicity and civil war. They argue that grievances are too common to explain

variations in civil conflict between nations (Fearon & Laitin 2003: 3). They rather adopt an opportunity approach to political conflict where they seek to analyses, as they say, the feasibility of insurgency. Weak governments, they argue, is what mainly improve the feasibility of insurgency. They find that ethnic fractionalization, measured by linguistic differences, has at best a marginal effect on civil war (Fearon & Laitin 2003: 3, 16).

Collier and Hoeffler (2004) whom have been mention before concerning state capacity conduct a similar analysis. They find very weak if any effects from their diversity variables (Collier & Hoeffler 2003:587-9).

Finally Cederman et al (2010b) create a dataset that takes into account the power relations between various ethnic groups in societies. Rather than measuring total fractionalization in populations this dataset gives information on the relative size of politically powerful and excluded ethnicities. They conduct an analysis of this data and civil war, and find that if large ethnic groups are excluded or if there is intense competition between groups in power, this is strongly associated with internal conflict (Cederman et al 2010b).

There is academic consensus that population size affects conflict positively, that is that countries with larger populations have more conflicts (Dixon 2009). However, it is not entirely clear how population affects conflict. It could be an indicator of opportunity as well as grievances; on the opportunity side a larger population could mean a larger recruitment basis for rebels, it could mean larger administrative challenges and thereby a weaker state (Enia et al 2011: 2636-7). As a grievance factor, a larger population could mean a larger degree of ethnic fractionalization. It could also mean an on average lager number of other divisions in society along, religious, class or ideological boundaries. Which of these explanations is correct is still in debate, but there is no doubt that population size is correlated with conflict.

2.6 General Theoretical Explanations of Internal Conflict

In the previous section, I looked at the research that inspired this thesis and some trends in recent conflict research. In this section, I will explore the two major theoretical explanations that seek to be applicable to all kinds of political conflict in the hope of generating hypothesis that can be applied to all conflict types analyzed here.

There have been theories of the characteristic of various political regimes and the nature of political conflicts for as long as there have been political regimes and political conflicts. The ancient Greek philosophers were the first to make a systematic analysis of various political regimes. Aristotle [Unknown] (1996) made the first classification system of political regimes. In this system regimes were classified by two criteria how many ruled and whether the regime was beneficial for the population at large or only for the rulers. In this system there where two kinds of public regimes the benign one was called polity and corrupted one democracy. The point being that at this time democracy was thought of as an unstable and conflict prone form of government. Aristotle and Plato both were concerned with the quality of leadership and the danger of despotic populists rising to power in a democracy (Aristotle [Unknown] 1996, Plato [Approx. 308bc] 2001).

The point is that the modern conception of democracy as a particularly peaceful form of government is quite new, and was developed through the reformation and Renascence, where particularly new interpretations of roman text on republican government along with the increased individualism of protestant Christianity. Through the enlightenment period these processes inspire and drive the development of liberal thinking which eventually end up in modern liberal democracy as we now know it. Machiavelli is the first theorist to view the internal conflicts of republican regimes as a source of strength (Machiavelli [1531] 1950). He belied that with the right political institutions such conflicts could be resolved peacefully before they were allowed to escalate, the energy that would normally be expended on these conflicts could then be put to advancement of the common good (Machiavelli [1531] 1950). Tocqueville ([1835] 2006) is the

first political thinker to take note of the importance of social moments and the differences that exist in them from state to state. In his analysis of democracy in America he makes the observation that social movement were far more common in America than in France and far more rarely ended in violence there (Tocqueville [1835] 2001, McAddam et al 1996: 45). In later times, and especially after World War 2 the amount of work done on internal conflict has skyrocketed. Reviewing all of it is both impossible within the bounds of this paper and unnecessary, this paper will therefore focus on the two main prevailing theories in the field, the theory of relative deprivation and the theory of political opportunity.

2.6.1 Relative Deprivation Theory

A staggering amount of work has been done to try to scientifically explain political conflict. Much of this work culminated in the relative deprivation theory of political conflict in the 1970s. This theory states that political conflict is a result of a discrepancy between what people have, both in materialistic as well as idealistic ways, and what they believe they are entitled to. Ted R. Gurrs book *Why Men Rebel* (1970) puts it like this:

> "(RD) is defined as actors' perception of discrepancy between their value expectations and their value capabilities. Value expectations are the goods and conditions of life to which people believe they are rightfully entitled. Value capabilities are the goods and conditions they think they are capable of getting and keeping. (Gurr 1970: 24)"

In this book Gurr (1970) builds this theory of relative deprivation and explores many assets of value capabilities, and under what conditions discrepancy between value capabilities and value expectations may lead to relative deprivation. First, I will review this theory's general explanation of political conflict, and then I will turn to its view on the role of regime in political conflicts.

The theory of relative deprivation builds on physiological theories of aggression as well as classical and modern theories of revolution. Gurr (1970) reviews work done up to that point in time and attempts to create a unified theory based on it. The core causality in relative deprivation theory is elegantly and intuitively simple, frustration leads to aggression. Particularly frustrations

that cannot be avoided or escaped lead to aggression against its source, in political circumstances such frustration is a result of extreme relative deprivation (Gurr [1970] 2011: 36)[1]. Gurr ([1970] 2011) goes on to deal with relative deprivation in detail. He makes the quite illustrative distinction between deprivation and disappointment, where the point is that disappointment is the result of unrealized aspirations while deprivation is the result of unrealized expectations. Therefore deprivation is much more likely to result in aggression precisely because it derives from expectations, things one takes for granted.

Gurr ([1970] 2011: 46) then considers how relative deprivation relates to a changing political and social world. He introduces a typology of three different situations that induce relative deprivation, all of which relate to the idea that people adjust their value expectations to meet their value capabilities. The first is 'decremental deprivation' where expectations are stable but capabilities decline. The second is 'aspirational deprivation' where the capabilities are stable but the expectations increase. Finally, there is 'progressive deprivation' where the expectations rise and the capabilities decline (Gurr [1970] 2011: 46). Progressive deprivation is probably the most severe type. The idea was introduced by Davies (1962) as the J-Curve hypothesis, and is common in societies that see steady improvements followed by a rapid decline. In such a society people begin to expect further improvements and thereby become progressively deprived when society regresses (Gurr [1970] 2011: 52). Relative deprivation is then the source of political conflict for Gurr ([1970] 2011), both the type and severity of political conflict can be explains by analyzing the sources of relative deprivation.

In regard to regime Gurr (1970) strongly emphasizes the importance of legitimacy relative to peoples value expectations. His theory is centered around explaining the motivations of dissidents in order to explain the occurrence of political conflict. This paper will later deal with theories that consider the impact regime has on the actual opportunity to rebel regardless of motivation, for now it is enough to note that RD only depicts one side of the coin of political conflict. Legitimacy for Gurr points to by which degree the people of a state feel that their regime is proper and deserving of support (Gurr 1970: 185). The argument is that people will

[1] [1] This is also *Why Men Rebel,* but a different edition. I cite them differently in case the page numbers are not the same.

become frustrated and take action if a regime does something that is considered outside their authority. Additionally people will typically direct this frustration first against the incumbents of the regime, then the institutions and finally the regime itself (Gurr 1970: 186). This is because people may be fine with institutional ways in which policy is created, even if they don't agree with the specific policy. Additionally people have “participatory values” I.e. an understanding of in what manner and to what degree they ought to be able to participate in society and politics. If regimes do not deliver on these expectations people we resort to increasingly, drastic measures to see those values are met.

Gurr (1970) also notes that in the current day (1970) the idea of legitimacy is tied to consent, where authoritarian regimes who maintain stability only through coercion immediately become unstable if the coercion stops. Gurr (1970:232) expands on the theme of coercion, introducing a concept of 'coercive balance'. Coercive balance is a concept that recognizes that the coercive capabilities of either a government or dissidents is meaningful only so far as it is considered in relation to that of its opponent. This observation is interesting as Gurr’s (1970) theory is primarily interested in the motivations or grievances of dissidents and not the government in explaining dissidence. However here Gurr (1970) touches on 'the other side of the coin' where the capability or opportunity to dissent is considered. What is more Gurr ([1970]2011:235) touches on what is the conclusion of Buahaug (2006) and the inspiration of this paper, that the type as well as the intensity of political violence depends on the balance of coercive capabilities between regime and dissidents. Gurr ([1970] 2011) then makes a few hypothesis concerning a small typology of political conflict. He hypothesizes that when dissidents are relatively weak small-scale violence is likely, such as terrorism or coups d’état. If balance is approximately equal then civil war is likely, and if either side is relatively strong conspiracies are likely.

In addition to the coercive balance, Gurr (1970) also speaks of an 'institutional balance. Institutional balance refers to the idea that the side who can offer institutional ways for people to achieve their value expectations gains legitimacy. In terms of a regime, this means offering institutional ways for people to achieve their material as well as ideological values. In a democracy, the main way dissent is avoided is by supplying legal and viable ways for people to achieve their ideological and participation values, while autocracies rely on coercion. Gurr

([1970] 2011) deals with the mechanism behind this in detail when he examines what decides the intensity and scope of relative deprivation. Amongst other things, the amount of opportunities that an actor has determines the scope and intensity of relative deprivation. Concerning regime it is mainly a matter of political opportunities, while other opportunities such as economical or even personal could be affected by regime as well. The purpose of democracy here is to provide legal, peaceful and viable political opportunities in order to alleviate or prohibit relative deprivation, while autocracies use repression in order to discourage people from acting on their deprivation. Similarly, dissidents can also gain legitimacy and support by offering institutions that a regime fail to provide. As an example this was done by The Muslim Brotherhood in Egypt during the Mubarak regime, where they provided welfare goods such as schools, medical attention, and clothes where the regime failed to do so, and thus gained legitimacy with the population (Walsh 2003).

2.6.3 Political Opportunity Theory.

Also known as, 'political process theory' this theory tries to explain political conflicts by investigating the political structure and process of a given political unit to account for what kind of opportunity it gives to dissidents. The core argument being that actors will always choose the path of the least resistance, as political dissidence is not seen as something that has a value in it shelf, but as something that is done only as a means to an end. In relation to relative deprivation theory, political opportunity theory concerns itself not with the motivation of dissidents but with what opportunity there is for rebellion regardless. The tradition bears strong ties and similarities with rational choice theory (Enia et al 2010).

McAddam et al (1996) review the literature that existed in this field up until then, and attempts to create a foundation for a unified political opportunity theory. For this purpose, they single out three main aspects of the theory.

The opportunity structure is the sum of all factors that influences the opportunity to dissent,

considered as a unified system (McAdam et al 1996). There are a large number of factors in any such structure and the factors could vary from country to country. However, in most cases such things as the nature of the constitution, the legal system, political institutions and the balance of power between dissidents and incumbents will have a large impact on the opportunity to dissent.

Mobilization Structures account for the ability of dissidents to organize, as well as the general ability to cooperate in a society. The opportunity for dissent is considered greater in societies that have a culture that emphasizes cooperation and where people are experienced at organizing themselves. The logic is quite simply that in order to have any kind of dissension you have to be able to organize it. Extending this logic, what degree of organization a group is capable of may influence what sort of dissension they choose (Tilly 1978, McAdam et al 1996). The ability and freedom to organize is something that is best assessed qualitatively as it is a quite complex phenomenon. However, it seems reasonable that those features of dissidents should be more prominent in consolidated democratic regimes, as democracies encourage cooperation and allows freedom of organization.

Framing refers to the physiological idea of cognitive frames. Frames can be thought of as bridging the gap between what one should expect from political opportunity theory and what one can actually observe in the real world. Frames identify the perception actors have of their political opportunity structure. Because even though there may be a good opportunity for dissent this may not be clear to actors and they may also overestimate their chances and try dissent in a hopeless situations (McAdam et al 1996). In other words it accounts for actors not having perfect information.

Tarrow (1996) then goes on, in a sub chapter of McAdam et al (1996), to consider what opportunity structures can be discerned at the state level. The relationship he describes is curvilinear, where larger political opportunity is not associated with more, or with more violent protest, but where a mix of restriction and opportunity is associated with both. At the same time either extreme opportunity, or extreme lack thereof, is associated with less protest. This makes sense in light of the finding that mixed regimes are most associated with dissidence.

Like Gurr (1970) McAdam et al (1996) see legitimacy as key in regard to regimes. As Oberschall (1996) puts it: "lack of regime legitimacy is an opportunity for opponents (Oberschall 1996: 95)." In Gurrs (1970) theory regime affected relative deprivation. In political opportunity theory regime is a central part of the political structure that determines what types of political engagement that is possible and, opportune. The point of democracies providing legal peaceful and viable forms of political engagement is not here to alleviate relative deprivation, but to make the relative cost of other, potentially violent forms of engagement higher. Authoritarian regimes by comparison use repression in a larger degree to make the cost of any kind of unwanted political engagement high. A good example of this theory is an analysis of transitions form authoritarian regimes to democracies done by Linz & Stepan (1996). In their model the type of regime is critical in determining what manner of transitions may take place, as different regimes open different opportunities for the opposition. In other words the opportunity structure determines the most likely form of political activity, because actors evaluate political activities on a cost benefit basis, and regime is a central part of that structure.

Using the fall of The Soviet Union as an example Oberschall (1996) goes on to make a classic point that was also mentioned by Gurr (1970), about authoritarian regimes and dissidence. That an authoritarian regime that begins to open up more opportunities for political participation will quickly erode their authority, and frequently suffer rapid collapses. If dissidents can frame the situation in such a way that the regime seems immoral or illegitimate then they can erode its authority, and contrary if the regime can frame the dissidents as terrorists, criminals or similar, they erode the authority and legitimacy of the dissidents.

2.6.4 The Greed Grievance Debate.

In the introduction to the fortieth anniversary edition of *Why Men Rebel* Gurr ([1970] 2011) takes the opportunity to assess the applicability of his theory of rebellion in the current world. He makes several points about where *Why Men Rebel* is inadequate today as well as some comments that are relevant to the greed versus grievances debate; I will focus on these here, as not all his

comments are relevant to the paper. First he maintains that understanding peoples grievances is still the first step to understanding political conflict, and it is not enough to use political or economic structures as explanations, but one must understand how people view them (Gurr [1970] 2011: X). However, he concedes Tillys (1978) view that mobilization structures are also important in understanding political conflict, and calls for a synthesis between his and Tillys (1978) models. He later makes a point that the assumption that was made in *Why Men Rebel* that people who react to relative deprivation act irrationally was a mistake (Gurr [1970] 2011: XII). This does not mean that he thinks that the rational choice theory is right either, rather he expresses that neither of the theories are correct, as the both assume too much about rationality *a priori*. Finally, Gurr ([1970] 2011) makes a point about the role of the state in analysis of political conflict, where he responds to critique that his theory does not adequately account for the state. First, he claims that the state is part of his theory, in that it can contribute to grievances, and holds that legitimacy is a verified explanation for political conflict. However, he allows that the question of why some governments use repression while others do was not answered in *Why Men Rebel* (Gurr [1970] 2011: XIII).

Much of the most recent research on civil war builds on work done by Collier & Hoeffler (1998, 2004). They argue for what is known as the supply side of civil war where the idea is that civil wars occur where there is an opportunity for them to be successful (supply), as there is always some miss-contented group or other in any society (demand) (Collier et al 2003). In other words there is always demand for civil war but not always supply, and thus it is the supply side of conflict onset that is decisive in which countries have civil wars and not. According to political opportunity theory, the same logic should also apply to non-violent conflicts, but with different patterns. They separated opportunity theory and relative deprivation theory into the dichotomy of greed versus grievances, and started a debate about which of these theories best explained civil war.

Collier et al (2003) examine what has been done of research up until that point and conclude that neither greed nor grievances can explain civil war onset alone, however *every* society has grievances while civil war is relatively rare (Collier et al 2003: 89). Therefore predicting civil war is very difficult. In order for a conflict over some grievance to become violent there need be economical incentive, and a weak or illegitimate enough government for violence to be a viable

political means. Collier et al (2003) consider economic development to be the factor that best explains civil war onset, as this correlates with conflict regardless of other factors. Looking back at the above-mentioned theory of political opportunity and of relative deprivation it seems that it is opportunity that decides which conflicts become violent. As Collier et al (2003) put it: "[...] no military and financially *viable* [emphasis added] opportunity to promote a political agenda by rebellion will be missed (Collier et al 2003:89)."

Enia et al (2010: 2632) wrote a review article summarizing findings in the debate. They claim that the grievances theory lacks empirical support. They do however emphasize that some very recent studies show limited support for the theory, and that it is solid and intuitive on the abstract plane, despite lack of empirical support. As for the theory of political opportunity, Enia et al (2010) claim that there is more empirical support for this theory. In particular, they point out that economic development and population size has been shown to be very strongly associated with civil war. Enia et al (2010) then go on to criticize the current theoretical dichotomy in the field. Their main criticism is that the dichotomy is neither mutually exclusive nor useful as an analytical tool. The criticism that is leveled at both theories is that they lack an understanding of the causal mechanism that leads from either the presence of grievances or opportunities to actual dissent. There are lots of countries that have grievances that don't experience dissent, similarly there are lots of countries where there are great opportunities for dissent that is never acted on. Towards the grievances theory specifically they claim that it fails to account for any other motivations. Then they claim that attributing indicators like GDP, population, state capacity and regime entirely to either opportunities or grievances is less than ideal (Enia et al 2010). The reasoning they give for this is that several of these variables could be catching effects of both theories. One of the examples that they give of this is that lack of economic development can both be a source of opportunity through decreasing the cost of dissidence relative to inaction, and a course of relative deprivation compared to better times or other people. Additionally the two theories may explain the same conflict at different times, for example, a grievance-based conflict founded by diamonds may evolve to become mainly economically motivated as more and more economically motivated people join the rebels. Finally, it is a problem of both theories that they don't adequately account for the role of the state but focus mainly on the opportunities or

grievances of dissidents.

If Enia et al (2010) are right about this then it is a major problem. Their solution to how it can be solved is that the two theories must both be considered as mutually necessary in order to explain dissidence. They also argue that the specter of grievances ought to be opened up in an attempt to account for other motivations for dissent. To this end they champion the idea of using the term willingness, which is meant to encompass all sorts of motivation rather than only grievances.

In summary conflict can be explained in terms of both frustration (grievances) and opportunity (greed). As frustration is constant one looks for opportunity in order to explain conflict onset. However if one wishes to understand a specific conflict one must understand the frustration it is based on. It seems to me that the opportunity approach is most helpful in a statistical investigation, as representing grievances a way that makes sense statistically is very difficult. Particularly if one wants to represent the individual grievances and histories behind every conflict, it becomes impossible to represent in a data matrix. Statistical analysis must rely on some amount of generalization in order to uncover general patterns. However, for case studies seeking to identify specific casual mechanisms, trying to understand people and their grievances seems central. If these two approaches can supplement each other perhaps they can form a cohesive and generally accepted theory of political conflict.

The relationship of regime and conflict is not entirely understood, and there is much discussion in the field on whether it has a separate effect or if it is better explains as a co variant with economic development and/or state capacity and stability. It seems to me that it is possible that effects found from regime in early statistical studies such as Hegre et al (2001) could in fact be cashing the effect of other variables because of the methodological problems involved in measuring regime and conflict respectively. As Gleditsch et al (2014) calls it, the increasing disaggregation of variables in peace research may help uncover the actual correlations. It seems quite unlikely to me that the form of regime, that is the way in which politics are conducted in a country, should have no impact on the forms of conflict found in that country independently of other factors. That it not to say that they do not affect each other, they probably do, but I would

argue that to rule out regime entirely seems counter intuitive. From a perspective of opportunity regime is central in explaining both what forms of dissidence that is available to actors and which of them ought to be successful. In terms of grievances regime ought to have some effect as different regimes deal with internal grievances in different ways, as well as through the effect of people comparing their political value capabilities to those of others thus creating deprivation relative to that group if their capabilities are lower.

2.7 Hypothesis

In relation to regime and conflict, what relationships are to be expected? A challenge for this paper is to create unified set of hypothesis based on research done on separate conflict types. As seen in the research section much of the work done in conflict studies naturally focuses on a single type of conflict. In order to attempt to create hypothesis about the relationship between regime and several types of conflict I looked at that try to explain political conflict in general (grievance and opportunity). Inherently this means that this paper simplifies the relationship, the effect of regime on each of this conflict types could easily be an entire paper, but the point is to use new data to facilitate a unified analysis of several conflict types, therefore a unified set of hypothesis is also required.

If all of the empirical evidence I considered it seems to me that there is a relationship between regime and conflict, as the majority of studies do find *some* relationship, even thou it is not entirely robust across different studies. However it seems that it is not entirely clear what the statistical relationship and the causal mechanisms that link regime *type* and regime *consolidation* and *stability* to each other is, and how much of conflict variation that can be explained by either.

In any case, it seems to me that protest actions ought to be associated with democracy. This because democracies repress small-scale violent dissent to a smaller degree than authoritarian regimes, and completely allow non-violent protest actions. Therefore, the opportunity for this sort of dissent is greater in democracies, and I will expect that: *H1: Democracy will have a positive effect on non-violent protests.*

Empirical evidence seems to tend towards severe internal conflicts being less common in democracies than other regimes, although there is some disagreement about whether, or to what degree, this is explained by stability or economic development rather than regime. As for theory; the idea of democracy as a peaceful means of conflict resolution seems to have a lot of support. Therefore I expect that: *H2: Democracy should affect more severe forms of political conflict negatively*.

The opening up of participation combined with repression should help to both create opportunities for dissidents and fuel grievances toward the regime. Therefore I will expect that: *H3: Autocracies should see more of any kind of conflict.*

The shear amount of repression in consolidated authoritarian regimes should restrict opportunity for any kind of participation. While some have found slightly higher correlations between authoritarian regimes and civil war than democracies there is in theory nothing that indicates that they should have more forms of violent conflict than democracy. As for non-violent conflict everything indicates that these ought to be far more common in democracies and authoritarian regimes. *H4: Authoritarian regimes should impact all forms of conflict negatively.*

3.0 Data and Methodology

In this section, I will discuss the data I have chosen to use in my analysis and how I use these to operationalize the various variables, the methodology of the analysis itself, and methodological problems involved with both.

The section will begin with discussing why a quantitative method was chosen, the advantages and challenges of this method. Then I will consider the data, first concerning the dependent variable and then concerning regime, then control variables, and finally move on to consider the analysis itself.

3.1 The Statistical Method

The method of this paper will be to conduct a statistical analysis of event counts on a country year or monadic/national level (Gleditsch & Hegre 1997). There are several reasons that a quantitative research method was chosen. Firstly there are several methodological reasons. This paper begins with a theoretical motivated hypothesis about the relationship between regime types and conflict types. Based on previous research and theory, discussed above, the paper hypothesises that the logic driving the observed difference between various kinds of civil war and regime extends to other types of conflict as well. In order to test this hypothesis a statistical test is assembled to see if the same relationship can be found for several types of conflict.

There are several advantages to the quantitative statistical approach. The goal of a quantitative study is to find general relationships. The most favoured method for doing this is the experimental. Experiments involve manipulating data in order to uncover general relationships by observing the effects of the manipulations, preferably in a controlled environment where all variables that affect the data are accounted for, so that the exact causal relationship is revealed (Moses & Knudsen 2007). However this is obviously neither particularly possible nor ethically defensible in the social sciences in general, and even more so in conflict studies. When seeking a

general pattern the method of choice in the discipline has therefore been the statistical method. The statistical method relies on analysing data on things that have already happened. Statistics attempt to infer relationships that cannot be directly observed in the data through analysis. In the case of this paper, it will be used to try to infer whether there is a relationship between regime and conflict that cannot be directly observed in the data.

Another advantage is that given replication data and exact information of what has been done with the data it will be possible for anyone to replicate the results of a study. Which means that the method and results of my work can be assessed evaluated directly, rather than being dependent on my subjective interpretations. See appendix F for a download link to replication data for this study.

Finally, the method is appropriate to the question at hand, as regimes are complex phenomenon that are practically speaking difficult to assess quantitatively. Statistical analysis allowed for the explorations of several kinds of regimes and several conflict types, while a qualitative project would be limited to one or a few cases.

Additionally there are some practical reasons for choosing a quantitative approach. Much of the research that has been done in this field previously has been conducted statistically using country years, which means that data both on regime and control variables is primarily available in that format. Additionally even if a qualitative method had been appropriate for the question at hand, gathering primary data of any kind is both resource and time consuming neither of which I have much.

On the other hand, there are weaknesses in the method. Qualitative research could identify causal mechanisms that related to specific phenomena, i.e. specific regimes or specific conflicts that might be missed by a broad sweeping quantitative study. The quantitative method involves a fair portion of simplification. For example, the categorization of regime into democracy, anocracy and authoritarian regimes does not capture the entire picture of various regimes. Similarly, events have to be put into manageable categories, which means that much data is lost (see below for a more detailed description of this). In other words, quantitative method runs the risk of failing to

see causal relationships particular to individual cases, as caused by all cases in a category. Another weakness of the approach is that it assumes that the social world can be understood objectively, while the researcher is in fact part of it (Moses & Knudsen 2007: 145).

3.2 Data on Conflict and Challenges with Event Data.

The data on conflict in this paper is taken form The Global Database of Events, Language and Tone (GDELT) which is an automatically coded event dataset made by Kalev Leetaru, Yahoo! and The Institute for The Study of Diplomacy at Georgetown University(Leetaru & Schrodt 2013)[2]. The introduction of large-scale event data has been seen as a possible big step forward for the social sciences, as it makes information on social events more directly available (King 2011, Gleditsch et al (2014). The use of these kinds of data in conflict research is a new phenomenon and a very promising one; however there are also certain challenges to overcome that will be discussed here.

Events are simply things that have happened, typical taken from news reports. Simply put, event data are lists where each event is written in a sentence that contains codes that give information on that event through the use of a code book. The GDELT data then exists as a series of .text documents where each line is a separate event. Because there are a large variety of event types in GDELT some selection is necessary, I will come back to the exact selection later.

The GDELT data is automatically coded based news reports form a number of sources. Because it is automatically coded, it has the simple advantage of other datasets of being very big. The total number of events in the dataset at the time of writing is just above 200 million, in the words of an unknown genius; 'quantity has a quality all of its own'. The computer analyses news reports

[2]Apparently Leetaru has been accused of stealing the text sources for GDELT, as far as I can tell, whether he is guilty or not, this has no impact on the scientific properties of the data-set (http://www.state.il.us/court/R23_Orders/AppellateCourt/2013/4thDistrict/4130290_R23.pdf).

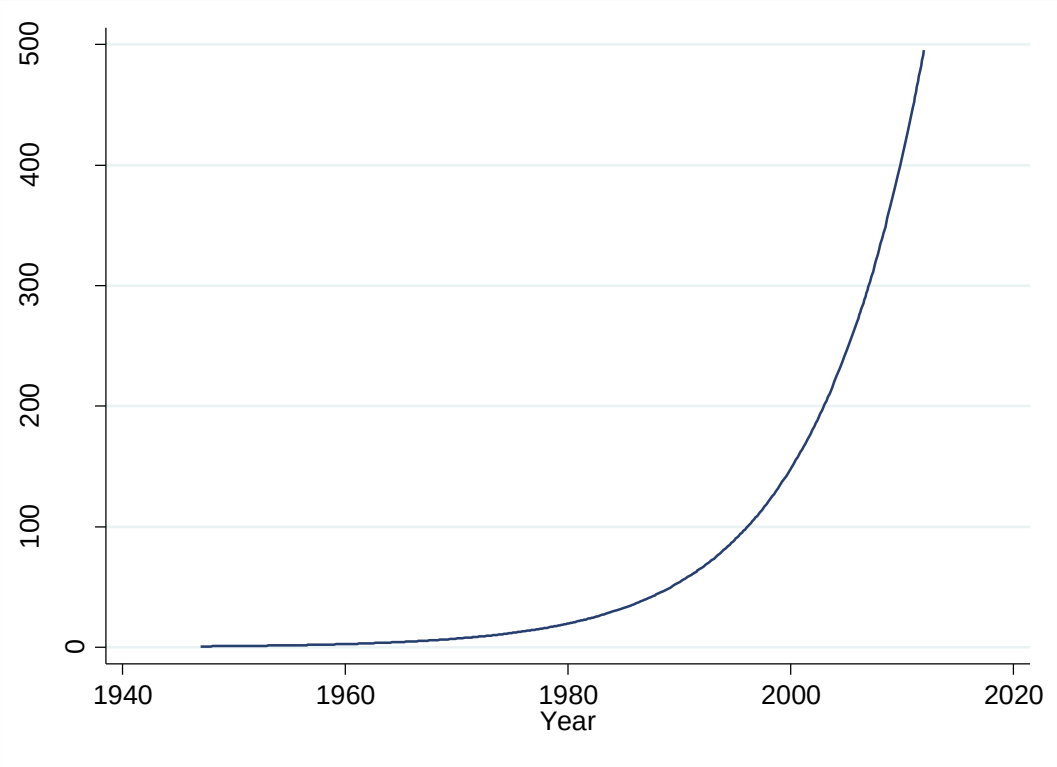
and writes the result of its analysis into a line in a text document. This line consists of a number of numerical and string variables that list a huge amount of information concerning the event in question. This is the basic premise of the data, however there are a number of questions that can be asked concerning the methodology of these data; how reliable is the computer interpretation of text? What sources are used? How is the data coded? Is there danger of duplicate reports? Etc.

In order to answer these questions we have to explore a couple of five letter acronyms, TABARI and CAMEO. Textual Analysis by Augmented Replacement Instructions (TABARI) is the actual software that codes the GDELT data from news reports. The inner technical workings of this program is highly complex, but the general idea is understandable. TABARI uses pattern recognition software to analyse short pieces of text such as headlines and wire services (Schrodt 2011). Essentially TABARI is a piece of software that is very good at sentence analysis. It looks at these short pieces of text and identifies (1) the actor (proper nouns), (2) verbs and (3) phrases that explain the meaning of the verbs (Schrodt 2011). However, it is important to note that TABARI is only good at this with short simple sentences, and makes mistakes when confronted with unusual grammar or complex sentences. This is because it is specifically coded to analyse news where the language is usually in the form of X does Y because Z. It is important to note though that TABARI is fed only news that is formatted to be readable by it, according to the author of the TABARI documentation (Schrodt 2011). The reason for this is that TABARI relies on what is called 'sparse-parsing' of sentences, which means it does not do a full analysis of the syntax, as some programs can do, but looks only for the three factors mentioned above (Schrodt 2011). The reasoning behind this is the desire for speed, the slimmer the program the more data can be gathered, and the faster the program will be in reanalysing data if the coding rules are changed. However there is a 'filter' of sorts in TABARI that separate very complex sentences which TABARI then skips, and which is later coded by a human. In addition, it is a problem that TABARI looks at sentences with complete disregard for context, so while it is quite clever it cannot be analytical like a human.

Furthermore, there is a bias in the data that is probably caused by the internet. There are simply far more events after circa the year 2000 when the internet becomes more common. This is due to there simply being far more news reports in the years where the internet is common, and that

data for the years before that are coded from archives while the data after the internet is partially coded live.

Figure 1: Predicted counts of events for all conflict variables over time.



The figure above illustrates the problem well. It depicts the predicted count for all conflict types by year. As year increases the predicted count increases. The data was obtained by running a simple count model with year as the only predictor. This model is not meant to serve the analysis it merely illustrates the general tendency towards more events later in the data. There could also be another reason for this general trend in the data. Populations has, as I mentioned in the theory chapter, been found to be a robust predictor of conflict (Hegre & Sambanis 2006). The population of the world has increased from 4.5 to 7 billion (US Census Bureau 2014), which probably means that the total number of violent events has increased as well.

However, it is not an option to disregard the later data, as some of the most valuable data comes from this period. In order to account for this effect a variable was constructed to represent the

increase in data as year increases. This variable, let us call it internet bias, is given as: internet bias = year – 1979. Where 1979 is the first year in the data and year is the time variable for country years. Thus the variable increases in value the newer the data is, if this bias has an effect on the analysis time should be significant and positive, because the newer the data the more events. Hopefully this variable will isolate some of the effect of this tendency.

A final problem with the data is considered by Quan Li (2005) in an article that examines the relationship of regime and terrorism. The problem is that data coded from news reports has a general bias towards more events of any kind happening in democracies, more so than should be expected from theory. The reason for this is differences in freedom of press. While democracies have full freedom of the press, anocracies and authoritarian regimes do not. Which again means there will be more events recorded in democracies than other regimes, not because democracies necessarily have more of these events, but because they are more frequently reported on in democracies. In other words, differences in freedom of press creates problems with data reliability. In addition, there could be similar problems in the bias of various news agencies, who are after all businesses and manly interested in whether something sells rather than if it is scientifically interesting. GDELT as well as most automatically coded event data tries to address this by using many sources, so that as many stories as possible are recorded, and as little of the individual news agency bias affects the data as possible. Similarly, press coverage could also be correlated with wealth. Poor and undeveloped countries may not have as good press coverage as rich countries.

On the other hand, there are several major advantages to machine coding; one is already mentioned, it is enormously fast. TABARI writes thousands of sentences every second! Coding at this level of magnitude is simply not possible manually. Not to mention that even if TABARI had been as slow as, in the words of the documentation author, “[...] legions of bored undergraduates (Schrodt 2011: 1)[...]” it would still be massively cheaper. Secondly, TABARI, unlike people, does not make mistakes. Critique may be levelled at the rules by which TABARI codes, but once those are in place it's consistency within those rules is 100% (Schrodt 2011). In addition to the advantage, that there are no coding errors there is neither any coding bias that cannot be accounted for by the coding rules, I.e any bias is reproducible. Additionally if the

sparse-parshing results in errors, they will be random, while human errors are frequently systematic (Schrodt 2011). Problems with random errors can be solved by statistical analysis, while systematic errors can be very difficult to correct. What is more if the coding scheme is changed and the entire dataset needs recoding, that can be done quite rapidly using computer coding. In one such instance they gained access to a computer used for scientific calculations, using this, TABARI coded 70'000 sentences a second of a total of 26 million sentences, which means that at the same speed all of GDELT could have been recoded in about 47 minutes (Schrodt 2011: 119).

Now, with TABRI in mind let us consider what actual codes TABARI writes. The codes written by TABARI are part of a system called Conflict and Mediation Event Observations (CAMEO). CAMEO is a system of number and string codes that distinguishes actors verbs and tones, and lots of information about them. In this paper, I am looking for specific types of events done by specific types of actors. I am looking for non-state actors and politically motivated conflicts. CAMEO codes actor type has a three-letter string variable and event-type codes as numeric codes following a simple structure of category and sub category. For instance 141 means demonstrations generally, 1411 means demonstrations for leadership change, 1412 means the same for policy change, there are several of these subcategories and they are typically the same for each major category (Schrodt 2011, 2012). In addition, the subcategories are then organized so that they increase with severity.

Now I will give my reasoning for selecting the actors and event types that I have selected, and then show how exactly I extracted the data from GDELT and how it was aggregated and what aggregation choices were made and why. To see the exact codes that were selected see the appendix A.

For actors all non-state, non-sivilian, actors were selected. The selection of actors is logical, as this paper is interested intrastate conflict. It makes sense to exclude governmental actors as governmental repression and one-sided violence is not part of this study. No further distinction is made between types of actors as this paper is trying to answer what happens in which kinds of

regime, not who does what in which kinds of regimes.

For events, however selection was a bit trickier. The main challenge was to deselect event forms that were not politically motivated, but of another, usually criminal, character or that were state specific. In brief, there are six categories of events in GDELT that have been used in my data, in which there are several sub-categories. These are:

1. Non-Violent Protest (peaceful demonstrations)
2. Violent Protest (demonstrations that turned violent)
3. Coerce (arrests, destruction of properties, and the like)
4. Assault (small scale, typically unprofessional, asymmetrical and mostly unarmed fighting, kidnappings torture, assassinations, suicide bombings etc).
5. Fight (conventional armed violence, typically of a military nature)
6. Unconventional mass violence (genocide, ethnic clearings and use of WMDs).

The first thing that should be noted about these categories is that verbal conflicts are not included. There are many categories in CAMEO as well as observations of this type of conflict in GDLET, however they are not the most relevant for intrastate studies, and it is problematic to consider many of these events conflicts. Secondly while there is a quad category variable the gives a rough impression of the event type in CAMEO, I thought it was too rough and I wanted to go into the sub categories to deselect event types that are outside this papers scope. As an example, I deselected any variation of the strike variable that did not state a political goal in order to avoid polluting the data with strikes over wages etc. In addition, I think that many of the categories, and especially the assault category, are composed of very different kinds of events and do not necessarily form a logical category. I therefore wanted to extract events at the sub-category level so that I could create my own categories in STATA later. These categories are:

1. Protests
2. Violent Protests

3. Small Scale Violence
4. Asymmetric Violence
5. Military Violence
6. Mass violence

The categories are ranked like in GDELT. They get progressive more difficult to organize and more severe, beginning with protests and ending up with mass violence.
The protest category contains various protest actions that were non-violent. Mainly selection was done on the strike variables as some of these were non-political, but directed towards wages.
The violent protest category is mostly the same except that the protest were or eventually turned violent.

The small-scale violence category is admittedly less consistent. The protest categories naturally belong together; this is less the case with the small-scale violence category. It consists of events that are more severe than protests, but perhaps less organized and more spontaneous, but more intentionally violent. Riots, blockades, and the like belong here. It is distinct from the asymmetric and military categories in that it only includes unarmed violence that requires far less organization and can be more spontaneous. Some may react to forms of violence such as sexual assault is included, as it is not typically politically motivated. However the actor types included in this data are. Therefore, even if not every incidence in this category is intended to have political consequences they do occur because of political conflicts.

The asymmetric category similarly contain events that are not necessarily governed by the same logic. What they have in common is that they are all forms of unconventional armed violence, such as bombings and targeting of civilians, in other words it is a category of terrorist tactics. The common factor among them is that they are all 'weapons of the weak', they take less resources and organization than the military category, but probably have the same goals.

The military category indicates civil war in this dataset as governmental actors are excluded, or at least incidents of conventional armed military violence. It covers all forms of conventional

armed warfare present in GDELT. Even though some of these like the use of air weapons and tanks is quite rare among rebels, there was no theoretical reason to exclude them or put them in another category. The data does however not accurately depict civil wars, but rather single instances of military violence, which do not necessarily only occur during civil wars.

The mass violence category contains events of massive unconventional violence, that is mass murder, ethnic cleansing and use of WMDs. Use of nuclear weapons have been excluded as there are (luckily) no events of non-governmental forces using them.

Now, why did I choose event counts on a country year level? Aggregating events to yearly counts by country loses a lot of detail on the events. Events that might be vastly different in significance are now clustered together simply because they are of the approximately same type. In other words, data on the severity of the event and or any symbolic meaning is lost. What is more, GDELT can identify the spatial location of events on a sub-national level. Aggregating the data on a national level loses potentially valuable data on where events occur. In a review of data trends in conflict research Gleditsch et al (2014) note how there has been trend towards disaggregating actors, events, strategies and resolution of conflict data. GDELT can potentially contribute in all of these categories of disaggregation, but this paper only disaggregates events. One place where GDELT could contribute to this in the future is the event categories for protests; these have different sub categories that allowed the researcher to distinguish different actor motivations. Along with distinctions of different actors, this could help to give a more complete picture of who participate in what manner of protest. However, the same is not true for the violent event categories, which is why this has not been attempted here.

I would however argue that the aggregation choices made here are sensible for a number of reasons. As Gleditsch et al (2014) stress, data aggregation choices must be made with the research question in mind. The country level is appropriate for the question at hand. Regimes are characteristics of countries and therefore analysing countries makes sense. Furthermore, because the vast majority of conflict research done before was done on a country year level, data on independent variables are frequently presented in that format which means it is practical to operate at this level, and makes it easier to put the results of this study into the context of

previous research.

The dataset I end up using was created with the aid of two `Python` scripts. For those not familiar with computer programming `Python` is an object oriented programming language[3]. The reason it was used is that it focuses on script readability, and has a straightforward logic to its syntax, which means it is relatively easy to learn. The first script simply extracts any sentence from the GDELT document if any of the chosen event-codes are present and the 'actor1typecode1' is of the chosen ones, that is, non-governmental. The second script creates a matrix in the random access memory (RAM) where it sorts the events by country year and lists the number of events in that country year. It then writes this into a tab delimitated text document that can be imported into statistics software. See appendix B for the complete scripts.

3.3 Sample

The data covers the years from 1980 to 2011 in 167 countries. There are 6493 potential observations taken from the GDELT data where 4697 are included in the analysis. The reason why these are different is that GDELT and the data on independent variables do not always cover the same countries or years and STATA automatically excludes and observations that do not contain data on both the independent and dependent variables. For a complete list of countries, see appendix C.

3.4 Data on Independent Variables

3.4.1 Regime

[3]Interested reader may want to know that GDELT has later published a tutorial on using Python with GDELT: http://nbviewer.ipython.org/urls/raw.githubusercontent.com/JamesPHoughton/Published_Blog_Scripts/master/GDELT%20Wrangler%20-%20Clean.ipynb and an online resource for extracting data from GDELT: http://analysis.gdeltproject.org/module-event-exporter.html

There are several sources of data on regime; Polity IV (Marshall, , & Jaggers, 2013), Freedom House (2013) and Vanhanen (2003). In this paper I have chosen to use Polity IV for three simple reasons; (1) they focus on institutional characteristic and seem less ideological than Freedom Houses rating, (2) they are based on more aspects of society than Vanhanen (2003) which is based solely on votes, and (3) it has good coverage and is available in a very easily access format. However, there is a major problem with the data that needs to be addressed briefly. Vreeland (2008) points out that two of the variables of that data, PARAREG and PARACOMP reference political violence between factions, which means that political conflict is endogenous to the Polity IV variable. This means that using the raw Polity IV data in an analysis of internal conflict could simply show that internal conflict is associated with internal conflict (Vreeland 2008).

Because of this problem I use Vreelands (2008) proposed x-polity index which is simply the Polity IV index minus the PARAREG and PARACOMP variables. I updated this index using the Polity IV data as the data provided by Vreeland only reaches to 2004. Unlike the original Polity index xpolity spans from -6 (most autocratic) to 7 (most democratic). Another reason that Polity IV is used instead of any alternatives is that Vreeland (2008) found similar problems with the other alternatives, but only Polity IV supplies the component data for their index, which means that the index can be reconstructed without the problematic components. In the analysis, the polity scale was recoded into the same dummy variables used by Vreeland (2008). Where democracy equals a score of 4 or higher on the xpolity scale, anocracy equals a score between -2 and 3, and authoritarian regimes equal a score of -3 or smaller.

The biggest problem with choosing independent variable data is to choose what variables need to be accounted for in addition to the primary bivariate relationship between regime and conflict. Any study of this kind walks a line between either excluding an important variable or including unnecessary variables. Schrodt (2014) writes a review article where he points out the most common methodological errors done in conflict research, where he emphasises the problem of having to many, and often correlated independent variables. Additionally another usual mistake is to use a dataset that has been analysed many times before. At least in this category this paper attempts to improve as quite few studies have used GDELT previously. In this study I use a

relatively small amount of control variables. The reasoning behind this is both that I want to avoid adding unnecessary variables and that I need controls that are relevant for a large number of conflicts and have been recently updated.

3.4.2 Wealth & Population

Wealth has been found to be one of the most important variables in explain conflict (Hegre and Sambanis 2006, Dixon 2009). This paper uses the updated v6.0 data on GDP from Gleditsch (2002). There aren't many problems to consider with this data GDP and GDP per capita are well known indicators of wealth and their problems are equally well know. The main problem is underreporting or miss reporting from various governments. As for the dataset itself, it is well referenced, the only problem being that the edition that is extended to 2011 is a beta version; however, both GDP and population are apparently complete. GDPPC is given as year 2000 $US.

3.4.3 State Capacity & State Stability

As was shown in the theory chapter state capacity, state stability and regime are variables where the exact mechanisms of how they affect conflict and how they affect each other's effect on conflict is not fully understood. It therefore seems reasonable that in order to isolate the effect of regime alone, it is necessary to account for the effect of stability and capacity. State capacity is a variable were there are few good data sources, as it is quite difficult to operationalize. Strand et al (2006) conduct an analysis of a database of irregular leader entry. They find that it is associated with civil war and that there are far more instances of irregular leader entry than of polity score change in the polity dataset and therefore conclude that polity is a poor indicator of stability. However, this dataset only reaches to 2000, which is a big reason against using it in this paper.

Hendrix (2010) reviews the various measures of state capacity that has been used in conflict research. The various approaches focus on military capacity, bureaucratic capacity, tax extractive

ability or quality of institutions. Hendrix makes a series of valid points regarding the validity and the ability of various measures to distinguish between different causal mechanisms. However, an overarching problem with the various data is that it does not have sufficient coverage for the purpose of this paper. The best possibility were GDP, relative political extraction and Polity squared. Among these I first favoured political extraction as it provides a theoretically sound and tangible measure of state capacity, the difference between actual tax extracted and theoretically possible tax extracted. However, the dataset provided Kruger & Tammen (2012, 2013) has serious problems with missing values and sub optimal country coding. Second polity squared is simply the Polity index squared, which gives a new variable ranging from 0 to 100 where, rather than the for of regime, the higher the score the more consistent the regime (Hendrix 2010). Similarly do to the problems with the Polity index discussed above Vreeland (2008) composed a similar scale disregarding the problematic variables in the Polity index. However, I am concerned that Polity squared is too similar to my regime variable which involves a risk of multicollinearity. Therefore, GDP remains the best possible measure of state capacity for this paper. In addition to being readily, available GDP has the advantage of being highly correlated with several measures of bureaucratic capacity (Hendrix 2010: 277). There are however several problems with this that the reader should be aware of when viewing the results of the analysis. First of all GDP can be both a cause and an effect of bureaucratic excellence. Secondly, it relates to conflict also as a measure of economic development, not just administrative capacity. Specifically GDP is considered by many to affect the cost/return logic of joining an insurgency thus affecting the opportunity structure (Collier & Hoeffler 2004, Hendrix 2010). The reader must therefore consider that the causal mechanism between GDP and conflict is not singular when reviewing the results of the analysis.

As for state stability I have chosen to use the same kind of variable as for example Hegre et al (2001) and Fjelde (2010) which is a decay variable of time since last regime change. The value of this variable decreases the longer it has been since a country experienced transition. The exact formula is 2^(-durable/3), where durable is the Polity IV variable that gives the years since last regime change, and 3 is the half time parameter. It makes sense to think that the longer a regime has lasted the more stable it is. I call the variable state instability, as a greater value equals greater instability.

An alternate and perhaps better measure of state stability exists in Scalar Index of Polities (SIP) dataset, which measures the inconsistency and stability of political regimes through tracking irregular leader entry (Gates et al 2006). However, again the data only extends to the year 2000.

3.4.4 Ethnicity

The importance of ethnicity as an explanatory variable is been subject to discussion for some time. However, it has been found to be robust for low scale conflicts for a long time (Hegre & Sambanis 2006). However, recent research has also found it to be significant in explaining civil war (Cederman et al 2009).

On ethnicity there are several options but I have chosen to go with GROWup Ethnic Power Relations data (Cederman et al 2010b, Hunziker 2013). The reason for choosing this data is that it gives information not only on the ethnically composition of a country, but also on the power relation between ethnic groups. That is the relative sizes of the groups in power vs the entire population as well as the same for groups that are excluded from power. This gives the possibility of accounting for the role of the state and nationalism in the relationship between ethnicity and political conflict, and is therefore in my view a step up from simply measuring diversity, as diversity alone does not necessarily lead to conflict. The GROWUP data contains many measures of ethnic phenomena. I have chosen to use the variable the indicates the size of the excluded population as a predictor of conflict as this variable should theoretically affect several kinds of conflict, while other variables in the dataset are more specific. The variable takes values between 0 and 1 where the value indicates the portion of the population of a country that is excluded relative to the total population (Hunziker 2013). In order to facilitate interpretation a categorical variable was assembled. It takes the form of a 10 point scale ranging from 0 to 9 where 0 indicates that zero % of the population is part of a politically excluded ethnic minority, 1 indicates that up to 10% are excluded, 2 that between 10% and 20 % are excluded, 3 that are between 20% and 30% are excluded etc. This makes it much easier to understand the effect of ethnicity in terms of incidence rate ratios, as they are given at one change in the value of the

independent variable. An important note on ethnicity is that different numbers of excluded people have been found to be linked to different conflict types (Cederman et al 2009). This means that I do not expect a linear relationship between this conflict variable and conflict.

3.5 Reliability and Validity

Above I discussed several problems with the sources and the method of gathering the data on conflict. As they are coded from news there are considerable problems with the reliability of the conflict data. For reasons of policy towards press freedom, differences in press coverage due to differences in wealth, as well as the effect of the internet are all likely problems in the data. This means that the data does not optimally represent the real nature of conflict, as there are considerable numbers of incidents that are never reported. However, it does not seem possible that these issues can be avoided as long as the coding relies on news reports, and any other source seems unlikely, at least for automatically coded data.

As for validity, there is little reason to suspect that the data is not valid. Wrongful reporting of conflict events is quite easy to expose, and not tolerated in journalist circles. Beyond this, any editorializing is impossible within the format that TABARI uses to code the data. As long as one is clear on the fact that GDELT lists single events of violence and not entire conflicts the data do measure what they are meant to measure.

3.6 The Regression Model

The analysis is a series of zero inflated negative binomial regressions. One regression model was fitted for each of the categories presented above, so that the differences in the effect of regime and control variables between them can be assessed. Zero inflated negative binomial regression

is a statistical method that is intended for use on count data that has excessive amounts of zero values, and is over dispersed. The inflation model assumes that independent variables have different effect on whether there are *any* above zero values, and how *many* above zero values there are. This aspect of the model must be stressed. The model assumes that a zero value can be the result of two different processes. In one proses a zero value is 'certain' or at least very likely, these are known as 'excess zeroes'. In the case of this analysis, an excess zero is the result of a country year where there is no opportunity or reason for conflict. In the second proses, a zero value is not 'certain', but occurs for different reasons. In the case of this analysis, a country year can have opportunities, but still not see conflict. The assumption is then, in terms of government and conflict, that variables can affect conflict in different ways. They can simultaneously have effects on the likelihood of excess zeroes, or initiation of conflict, and on the number of events, the scope of conflict. The zero inflated model conceits of two parts, a binary to model the excess zeroes and a count model to model the counts. In this case, the binary model is an ordinary logistic regression model, and the count model is a negative binomial model. The logistic model predicts the log likelihood of a value being an excess zero, while the negative binomial model predicts the number of counts. Therefore, the logistic model *cannot* be interpreted like a normal logistic model. It does *not* predict the odds of conflict, but the odds of a zero value being an excessive zero value. Therefore, if an independent variable, say GDP, has a negative coefficient then the log odds of conflict *increases* with GDP, and vice versa. In this analysis, for example, regime is seen to have first an effect on the probability of triggering conflict, then a separate impact on the probability of how many conflicts occur. The count model in this paper is a negative binomial model, which is an adaptation of a regular Poisson count model intended to account for over dispersion of the data. The regression expression is defined as:

$$P(Y_i = y_i \mid x_i) = \begin{cases} p_o + (1 - p_o)\left\{\dfrac{\alpha^{-1}}{\mu + \alpha^{-1}}\right\}^{\alpha^{-1}} & ; \ y_i = 0 \\ (1 - p_o)\dfrac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})}\left\{\dfrac{\alpha^{-1}}{\mu + \alpha^{-1}}\right\}^{\alpha^{-1}}\left\{\dfrac{\mu}{\mu + \alpha^{-1}}\right\}^{y_i} & ; \ y_i > 0 \end{cases}$$

As all the variables in the analysis are thought, in theory, to potentially affect conflict onset in both as a trigger and an escalating variable, they are therefore all included in both models.

In the theory section, I mentioned that Krain (1998) recreated the analysis of Powell (1982) because he wanted to apply more apt analysis tools to Powell's (1982) data. This serves as a good example of the utility of event count models. Powell (1982) used ordinary least squares regression to analyze event counts, which holds several problems. First OLS regression assumes negative values, which of course do not exist in event data. Second OLS assumes that a change in a variable from say 54 to 55 is the same as a change from 0 to 1 which with event data is unrealistic. For example, the difference between there having been one incidence of armed fighting and there having been none is arguably more significant than the difference between say three and four incidences. What is more, zero inflated event count models use probability distributions that account for the vast majority of observations having zero values. The data used in this paper is a typical example, each observation is a country year, and most have zero of any of the types of events that I am looking at.

Bagozzi (2011) makes an inquiry into the aptitude of zero inflated count models for predicting conflict. Using different conflict data, that was also coded with the TABARI program using CAMEO structure, he finds that the predictive capability of the zero inflated model is 12% better than that of a normal count model. He also finds that lagged conflict variables in zero inflated models improved the models significantly (Bagozzi 2011). Bagozzi (2011: 26) explains this by showing that the positive effect of recent conflict is exaggerated when models do not account for zero inflation.

Instead of a count model, it is quite common to define the dependent variables as dummy variables where conflict onset is coded 1, and then analyze this data using a logistic model. However, I have chosen to use a count model as I think it should be able to model the relationship of the variables more accurately.

3.7 Descriptive Statistics.

Table 1: Descriptive statistics of dependent variables

Variable	Observations	Mean	Standard deviation	Min	Max	Variance
Protests	*6493*	*22.5*	*148.3*	*0*	6087	22012.64
Violent Protests	*6493*	22.4	105.6	0	3391	11058.97
Small Scale Violence	*6493*	13.4	68.3	0	2443	8741.887
Asymmetric	*6493*	21.2	156.5	0	4151	18827.87
Military	*6493*	101.2	578.0	0	21291	334139.6
Unconventional Mass Violence	*6493*	0.37	2.4	0	168	8.823956

Table 2: Descriptive statistics of independent variables

Variable	Observations	Mean	Standard Deviation	Min	Max	Variance
Democracy	*10721*	0.774	0.417	0	1	0.174
Anocracy	10721	0.059	0.235	0	1	0.059
Authoritarian	10721	0.166	0.372	0	1	0.138
GDPPC	9626	8.342	1.234	4.888	13.357	1.522
Population(ln)	9626	8.460	2.129	1.791	14.096	4.533
Ethnicity	10721	1.48	2,168	0	9	4.700
State Stability	*5339*	0.229	0.334	4.27e-21	1	0.111

3.8 Assumptions of the Model & Other Challenges

Negative binomial regression assumes that the data are clustered around the lower values just like in Poisson regression. However, negative binomial regression does not follow the exact same probability function as Poisson regression, but assumes that the data are highly dispersed. What this means in praxis is that the variance of the dependent is expected to be greater than the mean. In my case it is for all variables.

Several of the variables have been changed for the analysis to deal with various methodological problems. First, the dependent variables have been time lagged in order to prevent problems where last year's conflict affects this year's conflict, creating serial correlation due to endogenous data. The time lag is of 1 year. Secondly GDPPC and Population has been given at their natural log because the effect of these variables are likely logarithmic. This also helps to facilitate interpretation of the model as the coefficient for one$ change would likely be minute and therefore difficult to interpret.

4.0 Analysis & Discussion

In this section of the paper the results of the analysis will be presented and discussed. As this paper seeks to analyze six different types of conflict a separate model was constructed for each. The first step shows some bivariate relationships between regime and conflict. For step 2 a zero inflated negative binomial regression model was constructed for each of the dependent variables. Apart from the dependent these models were all exactly the same, using every one of the control variables discussed in the methodology section. The zero inflated negative binomial regression model can be interpreted in a number of ways. First, I will consider the negative binomial and logistic regression coefficients, and significance levels. As a note, the negative binomial coefficients indicate the change in log count with one change in the independent variable. The logistic coefficients indicate the change in log odds *of being an excessive zero* with one change in the independent variable. Then I will move on to consider predicted counts from the negative binomial model and predicted probabilities of excess zeroes from the logistic model.
Finally, step 3 seeks to explain unexpected findings in step 2. This step explores the possibilities of methodological problems in the analysis, and tests for possible interaction and curvilinear effects as well as influential cases.

This section will begin with presenting the results of the analysis for the various models. Then discuss the results, and finally conclude.

Table 3: Regression Models for Step 1

	Protests		Violent Protests		Small Scale Violence		Asymmetric Violence		Military Violence		Massive Unconventional Violence	
	Model 1		Model 2		Model 3		Model 4		Model 5		Mdel1 6	
	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log
Democracy	0.473***	-20.901	0.489***	-21.526	1.004***	n/a	1.516***	-22.410	1.171***	-14.450	0.116	-0.992**
	(0.079)	(3079)	(0.097)	(9241.2)	(0.070)		(0.089)	(45373.5)	(0. .0632)	(300166.2)	(0. .193)	(0.380)
Anocracy	0.295**	-20.093	0.473***	-22.395	0.989***	n/a	1.580***	-22.832	1.135***	-14.808	0.474**	-18.186
	(0. 111)	(6430.8)	(0.097)	(20573.8)	(0.105)		(0.128)	(93137.3)	(0.095)	(645881.3)	(0.236)	(5916.4)
Constant	2.785	-1.437***	2.739***	-2.895***	1.708***	n/a	1.688 ***	-4.496*	3.564	-17.493	-0.818***	0. 008
	(0. 066)	(0.157)	(0.055)	(0.452)	(0.055)		(0.074)	(3.385)	(0.049)	(483.004)	(0.193)	(0. 305)
N	4822		4822		4822		4822		4822		4822	

Standard errors in parenthesis
*p<0.10, **p<0.05, ***p<0.01

4.1 Step 1

Step one shows some simple binary relationships between regime and conflict. In this step six such binary models were constructed one for each dependent variable.

I will discuss the results by the independent variables. Beginning with democracy, we see that the logistic coefficient for predicting excess zeroes is insignificant in table 3 model 1. The relationship is negative, meaning that democracy increases the chance of conflict initiation. Towards the counts the relationship is however positive, and significant at the 0.01 level. The coefficient is 0.47 meaning that the log count of protests is .047 higher in democracies than authoritarian regimes, which is the reference category. For violent protests, model 2, the same direction and significant persists. The effect of democracy is insignificant in the logistic part. The effect of democracy is slightly stronger in affecting violent protests counts. The protest categories support the hypothesis about the effect of democracy. Democracies see more protests in accordance with *H1,* and fewer violent protests in line with *H2.*

Moving on to small-scale violence in model 3, we see that the effect of democracy is positive, significant at the 0.01 level in predicting the number of counts. What is more, the positive effect of democracy on small-scale violence is stronger than its effect on protests. The recent wave of riots and similar events was mentioned as part of the motivation behind this paper. In that regard, it is interesting to see that the effect of democracy on this kind of violence is positive and quite strong. This finding does not support *H2,* that democracy will see fewer violent conflicts. However, it does not entirely erode it as these are still not very grave forms of violent conflict. The model predicting small-scale violence is a normal negative binomial model as the Vuong test showed that a zero inflated model is not an improvement on a standard model in this case. However, if we look at the effect of democracy in the more serious conflict categories, in model 5 and 6, we see that the effect of democracy is positive and significant towards the number of counts here as well. The effect is strongest for asymmetric violence at 1. 516, in model 4. In the prediction of excess zeroes all the relationships are insignificant. This is a quite unexpected relationship. From theory and previous research one would expect democracy to have a negative effect on these kind of conflict (Hegre et al 2001, Hegre & Sambanis 2006, Collier & Rhoner

2008). The model predicting massive unconventional violence is not significant, see appendix D.

Moving on we can judge the effect of democracy in relation to the effect of anocracy. Form theory we should expect anocracy to have a stronger positive effect on conflict than democracy. We find the expected relationship in most of the models. Democracy has a stronger positive effect towards conflict than anocracy. Democracy also has a stronger effect towards counts of small-scale violence, which is somewhat unexpected. However, anocracies have more counts of asymmetric, military and mass unconventional violence according to the model.

How can these unexpected findings be explained? It could be that the effect of regime is contingent on other variables. For example, the negative effect of democracy on conflict is often thought to be contingent on wealth (Collier & Hoeffler 2004, Hegre & Sambanis 2006, Gratzke 2007). This is the limit of a bivariate analysis. You may find relationships between the two variables that are in fact spurious, and an effect of another variable not accounted for. In step two the same models are constructed using all the control variables discussed in the methodology section, GDPPC, ethnicity, regime instability and internet bias.

4.2 Step 2

The regression coefficients for step two are reiterated in table 4. The first thing that should be noted about them is that the number of observations (N) is smaller in this step than for step one. The reason for this is that the coverage of the control variables is not as good as that of the dependent and explanatory variables.

Table A1 in the appendix shows the significance levels, Vuong test, Chibar^2 and alpha values of all the models. All models are significant. The Vuong test indicates whether a zero inflated model is a better fit than a normal negative binomial model. Alpha is the log of the dispersion parameter and Chibar2 indicates the probability that a zero inflated Poisson regression would be a better fir than a zero inflated negative binomial model. All tests indicate that the models are a good fit.

Table 4: Regression Models from Step 2

	Protest		Violent protest		Small-Scale Violence		Asymmetric Violence		Military Violence		Massive unconventional violence	
	Model 7		Model 8		Model 9		Model 10		Model 11		Model 12	
	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log
(Ln)Population	0.551***	-1.077***	0.519***	-1.314***	0.523***	-1.096***	0.550***	-1.349***	0.603***	-1.363***	0.124***	-0.702***
	(0.0169)	(0.077)	(0.0161)	(0.0952)	(0.0200)	(0.0732)	(0.0256)	(0.108)	(0.0195)	(0.119)	(0.0408)	(0.0767)
(Ln)GDPpc	0.433***	-0.369***	0.243***	-0.453***	0.0133	-0.325***	0.220***	-0.369***	0.00128	-0.705***	-0.186***	-0.0850
	(0.0237)	(0.081)	(0.0213)	(0.0965)	(0.0261)	(0.0842)	(0.0305)	(0.0960)	(0.0225)	(0.152)	(0.0713)	(0.0867)
Anocracy	0.0114	-0.167	0.238***	-1.866**	0.661***	-0.507	0.812***	-1.273**	0.767***	0.0822	0.326	-1.045***
	(0.0853)	(0.334)	(0.0729)	(0.795)	(0.0899)	(0.434)	(0.107)	(0.596)	(0.0838)	(0.705)	(0.229)	(0.310)
Democracy	0.123**	-0.454**	0.0557	0.593**	0.762***	0.755***	0.760***	0.542**	0.845***	0.927***	0.220	-0.334
	(0.0600)	(0.201)	(0.0541)	(0.268)	(0.0654)	(0.214)	(0.0771)	(0.270)	(0.0577)	(0.327)	(0.195)	(0.231)
Ethnicity	0.0956***	-0.00461	0.139***	0.126***	0.131***	-0.0422	0.214***	0.0994**	0.210***	0.155***	0.0961***	-0.0908**
	(0.0126)	(0.0374)	(0.0115)	(0.0435)	(0.0146)	(0.0406)	(0.0168)	(0.0476)	(0.0131)	(0.0548)	(0.0340)	(0.0403)
State Instability	0.809***	-1.069***	0.633***	-1.088**	0.677***	-0.859***	0.754***	-0.563	0.794***	-1.594***	-0.339	-1.103***
	(0.0871)	(0.344)	(0.0796)	(0.424)	(0.0950)	(0.325)	(0.114)	(0.424)	(0.0864)	(0.592)	(0.248)	(0.380)
Internet Bias	0.0925***	-0.119***	0.0863***	-0.0970***	0.0611***	-0.0869***	0.0944***	-0.0440***	0.0620***	-0.152***	0.0576***	-0.0346**
	(0.00281)	(0.0112)	(0.00250)	(0.0119)	(0.00315)	(0.00988)	(0.00353)	(0.0114)	(0.00278)	(0.0206)	(0.0120)	(0.0145)
Constant	-8.933***	13.21***	-6.571***	14.02***	-5.079***	12.04***	-7.814***	13.52***	-4.102***	15.59***	-1.060	9.706***
	(0.276)	(1.037)	(0.240)	(1.187)	(0.295)	(0.976)	(0.359)	(1.275)	(0.266)	(1.616)	(0.697)	(0.938)
N	4,697		4,697		4,697		4,697		4,697		4,697	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<

Beginning once again with the effect of democracy, we see that the coefficients in model 13 in step 2 is significant at the 0.05 level and that the effect is negative. A negative effect means that democracy decreases the odds that a zero value is an excessive zero, and thereby increases the chance of conflict initiation. The corresponding negative binomial coefficient shows that democracy is positive and significant, which means that democracies see a larger count of protests than authoritarian regimes. This is as expected in *H1*. In model 8 we see that effect of democracy on violent protest is positive and significant in predicting excess zeroes. The effect of democracy on the number of violent protest counts is not significant.

In model 9 small-scale violence is affected by democracy in a more peculiar way. The effect on the chance of an excess zero is positive and significant. In the count model, democracy also has a significant and positive effect. This difference in direction of effect is interesting. It means that democracies on the one hand affect that chance of there being any small-scale violence negatively, and the amount of small-scale violence positively. This is strange as we should in theory the relationships would seem most likely to have the same direction. However, this is not an indication that the zero inflated model is poorly fitted, as the zero inflated model assumes that this kind of relationship is possible.

Similarly, I find in model 10 that democracy has a positive and significant effect on both parts of the model. Which means that democracies have a larger probability of avoiding conflict, while they at the same time have a larger count of violent events when conflicts occur. This finding contradicts *H2,* that democracies will see fewer violent conflicts. That said, as I discussed in the theory section the findings on the relationship between democracy and asymmetric violence is not robust across different analyses. Also the theory points to both positive and negative effects of democracy, this find does not seem so strange in that light.

For military conflicts, I find a relationship that is strange. As with small-scale violence, democracy has different effects on excess zeroes and counts here. In model 11 the effect on the likelihood of excess zeroes is positive, as is the effect on the counts. The effect of democracy is therefore twofold here as with small-scale and asymmetric violence. What is more, democracies experience more counts of military violence than anocracies, which is unexpected and does not

support *H2,* or *H3*.

In predicting massive unconventional violence, the effect of democracy is negative towards excess zeroes and positive for counts, however the effect is not significant.

Comparing the effect of democracy with the effect of anocracy I find that anocracy does not have a statistically significant effect on protests. Model 8 finds that anocracies impact the likelihood of excess zeroes negatively, and the number of counts of violent protest positively. The effect is significant for both coefficients. As the effect of democracy is insignificant on counts and positive on excess zeroes, I find that anocracies see more violent protests than democracies. This is in line with my expectations that violent conflict are more rare in democracies, *H2,* and that anocracies generally see more violent conflict, *H3*.

Incidents of Small-scale violence are only affected by anocracy in the number of counts; the coefficient for excess zeroes is not significant. The effect on the counts is not as strong as that of democracy, Which I will return to in considering the incidence rate ratios below.
Asymmetric events are according to model 10 more common in anocracies than in authoritarian regimes. In predicting excess zeroes the effect is negative and in predicting counts it is positive. The effect in the count model is very similar in strength to that of democracy.
For military events anocracy is only a significant predictor of the number of counts, the effect is positive and stronger than the effect of democracy.
Finally, anocracy reduces the chance of avoiding military events. It also increases the predicted counts, but this effect is not statistically significant.

All in all the findings in step two are not in line with the expected relationships. Anocracy frequently has a smaller positive effect on violent forms of conflict than democracy. This means that violent conflict is found to be more common in democracies than anocracies. This erodes support for *H2,* and for *H3.* However step 2 does support *H4,* that authoritarian regimes should see less conflict counts than other regimes, as both the regime terms have positive effects.

Finally, considering the control variables we see from the logistic coefficient for GDPPC that it

is a significant predictor of excess zeroes for all conflict types but massive unconventional violence. From theory, we should expect GDPPC to affect most conflicts negatively; however, it affects the likelihood of excess zeroes negatively for all conflict types, and is only insignificant towards small-scale and massive unconventional violence. Towards the number of counts GDPPC it is significant and positive for protests, violent protests and asymmetric conflicts. From theory I do, as mentioned, expect GDPPC to affect conflicts negatively. The effect is in fact negative on massive unconventional violence, and the effect is statistically significant. For military and asymmetric events, the effect is not significant. It might be that the positive effect of GDP here is explained by a correlation between GDPPC and democracy, where rich western democracies drive the findings in protests. So the findings for GDPPC are not that unexpected as it is only positive for protests and asymmetric conflicts. Protests are not very serious conflicts and the findings in previous research that indicate a negative effect of GDP are mostly taken from studies of civil war. As to asymmetric events, there is no consensus on the effect of wealth.

Population on the other had taken on the expected relationship for all of the conflict types in the logistic part of the models. It is negative and significant in all logistic models, therefore increasing the chance of conflict. In the count model, population size has a corresponding positive and significant value in all models. This means that population size increases both the chance of initiation and the number of event counts. In the count model, its strongest effect is on military violence, although the effect is relatively even all over.
The state instability decay function has the same effect as population, which is also expected. It is negative and significant in all logistic models except for asymmetric violence for which it is not significant. This makes sense as having a stable regime cannot necessarily aid a state against terrorism. The decay function is also positive and significant for all count models in step 2, except massive unconventional violence for which it is insignificant.

In predicting excess zeroes ethnicity is significant and positive towards violent protests, asymmetric violence and military violence. It also has a significant negative effect on massive unconventional violence. Towards the number of counts ethnicity is significant and positive for every category. It is interesting that ethnicity has a dual effect on many variables, both increasing the probability of an excess zero and the expected number of counts.

Finally the control for internet bias works as expected and is negative and significant in all logistic models. In the count models it is also significant and positive in all models. This means that in general the number of events in the dataset increases the closer we get to the present.

In order to more clearly show the differences in effects between the variables in the step 2 models I will present a table of incidence rate ratios. A quick note on incidence rate ratios; what is refereed to as counts in event data is in mathematical terms ratios, i.e. the number of incidents in a given period of time. Incidence rate ratios are the ratio of incident rates between variables, in this case the ratio of event counts in country years (rates), between different variables. For example if democracy has an incidence rate ratio of 2 democracies have 2 times the counts of authoritarian regimes (the reference) as it is a dummy variable, all other rates held constant. If a continuous variable has an incidence rate ratio of 2 the number of counts increase by two for every level of that variable.

In order to facilitate reading I have recoded the variables into standardized variables where the mean is zero and 1 = one standard deviation above the mean, 2 equals two standard deviations above the means etc. This means the there is no need to consider the different coding of the various variables, so that their respective strength can be compared directly. I have not recoded the regime variables in this way, as they are already dummy variables. In the table, the values therefore now show the ratio of event counts by one standard deviation change in the independent variables, except for the regime variables that show the difference in event counts between democracy/anocracy and authoritarian regimes. Since the incidence rate ratios, unlike the coefficients, have a multiplicative effect, values below one indicate a negative effect.
It is also possible to interpret the strength of a relationship by the negative binomial coefficients, who indicate the change in the log of the count of a given country year given one increase in the independent variable. However, the increase in the log does not really tell us much as it is not a very intuitive measure.

Table 5: Standardized Incidence Rate Ratios for Step 2.

	Protests	Violent Protests	Small-Scale violence	Asymmetric Violence	Military Violence	Massive Unconventional Violence
Democracy	1.130*	1. 057	2.143*	2.137*	2.328*	1.246
Anocracy	1.011	1. 268*	1.936*	2.252*	2.153*	1.385
ZGDPPC	1.706*	1. 349*	1.016	1.311*	1. 001	0.795*
ZPopulation	3.234*	3. 020*	3.042*	3.225*	3.606*	1.302*
ZEthnicity	1. 230*	1.350*	1.328*	1.589*	1.575*	1.231*
ZState Instability	1.310*	1.235*	1.253*	1.287*	1.304*	0.892
ZInternet Bias	5.305*	1.090*	3.013*	5.488*	3.059*	2.826*
Constant	1. 223*	1.222*	1.166*	0.743*	6.401*	0.312

* P<0.05

In table 5 we can see the incidence rate ratios for the negative binomial regression models in step 2. Beginning with democracy again, we see that for all of the more severe forms of conflict where democracy is a significant predictor the effect stronger than the effect of anocracy, except for asymmetric effect where the ratio of anocracy is greater. The effect of democracy on protest is also stronger than the effect of anocracy, but this is expected from theory. If we compare the effect of democracy and anocracy the three models where they are both significant, we see that democracy is about 0.3 stronger for small-scale and military events, while anocracy is about 0.1 stronger for asymmetric events. The differences between the two variables are in other words not large. The most significant effect that we can take away from the findings in step 2 is that anocracies and democracies generally see many more conflict events than authoritarian regimes. The strongest effect of democracy is in military violence where democratic regimes have 2.328 times as many conflict events as authoritarian regimes. A positive effect of democracy on these very serious event types is quite unexpected.

All the significant relationships with anocracy are also positive. While the effect is weaker for most conflicts than the effect of democracies, anocracies still see considerably more of all kinds of conflict than authoritarian regimes. The strongest effect of anocracy is on asymmetric violence where the incidence rate ratio is 2.252. The findings for anocracy are for the most part not surprising when viewed by themselves. Anocracy does as predicted have more conflict events than the reference, which is authoritarian regimes. However, in light of the effects of democracy it seems weak. Most of the theory reviewed in earlier (Hegre et al 2001, Hegre & Sambanis 2006, Enia et al 2011) finds that anocracies are more prone to conflict than democracies. However most previous research has focused on one type of conflict (typically civil war) and has used datasets where one observation is one entire conflict and not single events of violence. Therefore, it is important to note that this does not say anything about which conflicts the various events are from.

Finally, the control variables one can now compare their strength directly. It seems that population is the strongest predictor of conflict along with the internet bias variable. Ethnicity shows a very even effect across most conflicts. GDPPC does not show the expected effect on counts. It is positive on protests which is unexpected, but difficult to imagine. It may be that protest actions are more common in higher GDPPC countries because many of these countries are democracies. The explanation could be given of the fact that asymmetric violence is also more common in higher GDPPC countries. GDPPC does not have a significant effect on small-scale and military violence, which is unexpected. Finally, the negative effect of GDP on massive unconventional violence is expected.

4.2.1 Predicted Probabilities and Predicted Counts

In order to go more in depth into the relationship between the independent variables and the dependent variables as well as to attempt to reveal any relationships between the independent variables I have used the regression models in step 2 to generate predicated probabilities and predicted counts. Just as you can use normal regression models to generate predicted values for the dependent, you can generate predicted probabilities of excess zeroes from the logistic, and

predicted counts from the negative binomial part of the zero inflated regression models. This makes it possible to show the predicted values of the dependent at different levels of the independent variables.

Table 6 shows predicted probabilities of an observation being an excess zero, other values held at means. This helps to illustrate the effect of regime on these conflict types.

Table 6 :Predicted Probabilities of Excess Zeroes From Step 2

	Protest	Violent Protest	Small Scale Violence	Asymmetric Violence	Military violence	Massive Unconventional Violence
Democracy	0.15*	0.13*	0.23*	0.20*	0.05*	0.72*
Anocracy	0.18	0.03*	0.12	0.09	0.04	0.57*
Authoritarian Regime**	0.35	0.19	0.25	0.23	0.11	0.81

All other variables at their means.
P>0.05 in the logistic model from table (for convince).
**Authoritarian Regimes given by democracy and anocracy at 0, all other values at means.
Note: remember that this is the predicted probability of an excess zero value.

The results in table 6 are quite interesting if one compares them to the logistic coefficients from step 2. Keeping in mind that these are the probability of zeroes not of conflict we can see the relationship between the various independent variables in the model and not just the individual effects on the dependent. It also allows one to assess the strength of the effects and not just the direction.

The predicted probabilities of democracy compared to those of anocracy are not in all cases as one would expect from the regression models. For protest, the chance of an excess zero is higher in anocracies than in democracies, which is as expected. The discrepancy between authoritarian and democratic regimes is large; democracies have a 15% chance of being an excess zero while authoritarian regimes have a 35% chance, more than twice as likely. Which supports both *H1* and *H4.* The relationship would also support *H3,* but the effect of anocracy is not significant.

For violent protests the relationship is opposite, the chance of an excess zero is 10% higher for democracy than for anocracy. The probability is even higher for authoritarian regimes, which is in line with both *H2, H3* and *H4*. It is interesting that the relationship is so different from the one found by the incidence rate ratios. However, the effect of anocracy in the logistic model is insignificant.

Small-scale violence also has the same relationship with regime, the greatest probability of an excess zero value is authoritarian regimes, followed by democracy and finally anocracy. Meaning the greatest chance of small-scale violence lies in anocracies, and the smallest in authoritarian regimes. However, the relationship with anocracy is insignificant.

For military conflicts, the relationship is the same, but the difference between democracy and anocracy is much smaller, 0.01. Once again, the relationship with anocracy is not significant however. This is interesting in light of the unexpected findings from the count model. Democracy was found increase the likely number of military events, but has the opposite effect in the logistic model, although the difference between democracy and anocracy is not large and there is still a considerable gap to authoritarian regimes.

The same relationship also repeats for massive unconventional violence where that chance of an excess zero is 15% greater for democracies than for anocracies, and 24% greater for authoritarian regimes. This gives support to *H2, H3* and *H4,* as the relationship is significant and in line with what was predicted in the hypothesis.

4.2.2 Predicted Counts

Using the `margins` command in STATA one can obtain the predicted counts of a negative binomial regression at various values of independent variables. Table 7 displays the predicted counts of the dependent variables for democracies, anocracies and authoritarian regimes. All other variables where set to their means. The purpose of this table is to make clear the difference

in effect between the three forms of regime; as such, the total count is not as interesting as the difference in courts between the regimes. However as the other variables are at their means if there are interaction effects between any of them and regime this could affect the predicted counts. The margin command also sets the probability of an excess zero at it mean. It is therefore wise to keep the incidence rate ratios in mind when viewing the predicted counts, as the predicted counts show the effect of regime for an average country year rather than by it self. In addition to the incidence rate ratios, predicted counts allow us to see the effect of one variable at different levels of other variables. In the table, all variables are given at their means. I will first consider the effects in the table then move on to test the effects of regime at different levels of other variables in order to try to illustrate potential co variance between the independent variables.

Table 7: Predicted Counts From Step 2

	Protest	Violent Protest	Small Scale Violence	Asymmetric Violence	Military violence	Massive Unconventional Violence
Democracy	5.22*	7.50	5.79*	5.93*	43.74*	0.17
Anocracy	4.45	9.50*	5.93*	6.94*	40.72*	0.32
Authoritarian Regime	4.40	7.28	2.96	2.91	18.91	0.10

All other variables at means.

*P>0.05 in the negative binomial model.

Authoritarian Regimes given by democracy and anocracy at 0, all other values at means.
Predicted Probabilities of Excess zeroes at means.

The expected counts for protests is relatively uniform with democracy expecting almost one more count. This finding supports *H1& H4* but none of the other hypothesis, as only the effect of democracy is significant.

For violent protests, however we see that anocracy has a slightly higher expected count than democracy, while authoritarian regimes have an even smaller count. This is exactly in line with the predictions of the hypothesises, anocracies see more protest that turn violent than other regimes.

Small-scale violence shows a more expected difference between the regime types; authoritarian regimes have a much smaller predicted count than democracy and anocracy. That authoritarian regimes have fewer of these kind of events is expected and supports *H4.* In regard to asymmetric violence, the difference is also unexpected in regard to theory.

Asymmetric events seem to be more common in anocracies according to the model. Anocracies about 1 more predicted count than democracies. Once again authoritarian regimes have the fewest counts.

Military violence shows the most unexpected number of counts. First of all, all three counts are incredibly high. Second, democracy has a higher count, at 43, than the other regime types, anocracy 40, and authoritarian 18. The fact that democracies have a larger predicted count of military conflicts is in itself strange. The most likely explanation is that there are simply very many counts of military violence in the data compared to other variables. If one recalls the descriptive statistics the mean of the dependent variables are all around 15 except military conflict, which is 77.

Possibly this could be due to the press bias in the data discussed earlier. Military conflicts are bigger news than most other conflict types; therefore, they may be overrepresented in the data. Alternately, since all other values are given at their means, one of the control variables could have a curvilinear relationship with military conflicts. If this is the case it could have the greatest effect at a mean value and drive the results, as the counts are generally very large not just for democracy. Finally, it is important to remember that these are not 43 counts of civil wars but 43 counts of military violence. As an example of what an actual civil war would look like in the data, the largest count of this category is Syria 2013(2012 because of time lag) with 21 thousand counts of armed violence. Syria serves a good example, as it is one count of civil war and 21 thousand of armed conflict. It is also a conflict that has been given a lot of media coverage, and which is contemporary. In comparison, the Nicaraguan civil war took place at the start of my dataset from 1979 to 1991. There are just under 2500 reports of military violence in Nicaragua for all of those years in the dataset. Even though the Nicaraguan civil war was not as severe as the Syrian one is, this still illustrates the difference in reporting.

Finally, Massive unconventional violence shows the expected distribution of predicted scores. Anocracies have the largest score followed by democracy, then authoritarian regimes. The scores for this category are also by far the smallest in general, which makes sense, as these events are (luckily) quite rare.

In summery the expected relationship was found for protests, violent protest, asymmetric and massive unconventional violence. While small scale, and military violence shows unpredicted relationships. Regime does not seem to explain massive unconventional violence where both regime variables are insignificant.

I find support for *H1* In step 2 democracies have a higher effect on protest and decrease the chance of an excess zero value.

Many of the relationships in step 2 are unexpected. The relationship found in the logistic parts of the models are very different form the one in the count part. The logistic relationship is far more expected from theory, while the positive effect of democracy on military, small-scale and asymmetric violence in the count model is unexpected. It is manly the findings concerning military violence that are difficult to explain. As democracies seem to have more excess zero values for the most violent conflict types and is an insignificant predictor of violent protest it is not unexpected that it should have a positive effect in the count model. This merely indicates that democracies see more counts than the reference, which is authoritarian regimes. However, the effect of democracy on the counts of military violence is stronger than anocracy as well, which is more difficult to explain. The same is true of small-scale violence, but the violence types included in this category are not as organized, grave or generally detrimental to society as the asymmetric and military categories. That said the finding is still contrary to what I expected from theory. *H2* Stated that democracies should see fewer violent conflicts than other regimes. This is true according to the prediction of excess zeroes, which means that the chance of never seeing any conflict at all is larger in democracies than other regimes. I must be mentioned though that for military conflict non of the regime types had large probabilities of never seeing conflict. However, democracies that are not in the excess zero category see more events of violent conflict than other regimes. Therefore, support for *H2* is mixed taking all the models in step two in view. Another possible explanation of the generally large counts of military violence, and the

unexpected effect of democracy towards it, could be that there are influential outliers that are driving the results. In this case democracies that have unusually high counts of military violence. I looked at the data and found that there are certain democracies that have very high counts of military violence. They include mainly the countries Israel, Pakistan and the US. Pakistan has a very high count of 11 thousand, which could be a result of the war against the Taliban in Afghanistan spilling over the border, also the US drone program. More worryingly, The USA have amongst the highest counts of military violence in the data set in 2011. Other western democracies such as the UK and France also have worryingly high counts. This means that there are problems with the data for these countries. One probable explanation is that events that have taken place in Iraq or Afghanistan have been wrongfully coded as an events on US, UK or French soil. A similar pattern exists for the data on small-scale violence where the largest counts are also India, Pakistan, Israel and The US in the last four years.

Another possibility is that armed gang violence has been wrongfully coded as military. As civilians have been sorted out from my data in the `Python` script this is not due to my coding of the variables and must have happened in the coding of the original data. I also used the location code tied to the action and not the actors specifically to avoid actions taken by an actor in another country appearing as an event in their country of origin. In addition to the effect of democracy, the positive effect of GDP on conflict is unexpected. I have already speculated that as some research has shown (Collier & Rhoner 2008, Dixon 2009) the negative effect of both these variables may exist only if they work in conjunction. The unexpected results of democracy could also very well be due to the nature of press reporting and the freedom of speech. What is more a quick google search of internet coverage will show anyone that the majority of internet users live in the west, which means that there could be a colinearity between democracy and the internet bias as well. Add this to the assumption that press coverage is also correlated with wealth and that many democracies are wealthy and there could be systemic bias towards democracies in the data.

Another possibility is that the unexpected effect of democracy is in fact an unexpected effect for anocracy. Anocracies are mainly thought to be prone to conflict because they are not consolidated regimes and thereby unstable. Could it be that the relatively weaker effect of

anocracy on conflict is a result of anocracies only seeing more conflicts when they are unstable? *H3* is also partially supported. While democracies do see more counts of military and small-scale violence, they also have a higher chance of excess zero values for these categories than anocracies. This means that on the one hand democracies will see fewer country years with this conflict type, but more events in the country years that do have them. Anocracies however have fewer excess zero values and more counts of violent protest and asymmetric violence, therefore evidence for *H3* is mixed.

H4 Is supported for every conflict type. Authoritarian regimes see fewer counts and more excess zero values for all conflict types.

In regard to the control variables the positive effect of GDP has already been mentioned, but also the dual effect of the ethnicity scale is unexpected. It is frequently positive and significant in both models. This is true of democracy in some cases too. Since the logistic model shows the log odds of being an excess zero that means that, a positive effect indicates a negative effect on conflict. When the variable is positive in both models this then means that it has the opposite effect in the two parts of the models. This does not indicate that the model is poorly fitted, as it assumes that the effect of independent variables will be different in each model, but how can this effect be explained? Why should the log odds of an excessive zero increase the more people are excluded from politics on ethnic grounds? There is a strong possibility that the effect of ethnicity is curvilinear and that the effect is strongest at different levels of excluded populations for different conflict types. This is because the relative strength of and excluded ethnic minority can influence their goals and strategy, so that for example small minorities may choose asymmetric approaches while large may choose military (Buhaug 2006, Cederman et al 2009).

4.3 Step 3

In order to attempt to explain some of the unexpected findings from step 2 I construct a final set of models where I test for the effects suggested above.

First, I test for an interaction effect between democracy and GDP in all models. The reason for this is that democracy has been found in previous research to only have a negative effect on conflict in conjunction with wealth (Collier & Rhoner 2008, Dixon 2009). If this effect exist in my data as well, we should see a positive and significant interaction term between democracy and GDP.

In order to test for the effect of outliers in the data on small-scale military violence I construct a dummy variable that equals 1 if a country year has more than 4000 counts of military violence, and a a variable that = 1 if a county year has more than 500 counts of small-scale violence. There are in total only 8 observations of this scale all form Israel, Pakistan and The US, and in the period from 2008 to 2011, and all democracies. The reason that these variables were set to these values is that at above 4000 and 500 events respectively there are no other countries represented in the data and very few observations. These values then represent a break in the data between the cluster of observations around the mean and the extreme values of these countries. If the variable contributes to the high count for democracy we should see a significant and positive effect in the negative binomial part of the model. Most likely they will be significant but their strength should tell me something about their total effect.

I also test for the possible curvilinear effect of ethnicity. I test for this by including an intercept term in the models where ethnicity has a dual effect. This term equals the ethnicity variable squared.

In order to test if there is a reporting bias directed towards democracies I create a second interaction term between democracy and the internet bias variable to see if democracies have a larger increase in events by year than other regimes. I include this term in all models.

Finally, I test to see if there is an interaction between instability and anocracies in all models.

Table 8 displays the results of model 3 where interactions between regime and instability, democracy and GDP and democracy and internet bias is tested. I will first deal with the effect of the new terms included in step 3 in all models, then I will deal with the models that showed unexpected results in step 2 more in depth.

Beginning with democracy and its interaction term with GDPPC. Either democracy or the interaction term is now insignificant towards counts in every model. It is significant towards explaining excessive zeroes in the models for protest (model 13), small-scale (model 15), asymmetric (model 16), and military violence (model 17). In all of the models, the base term for democracy is now negative, while the interaction with GDPPC is positive. This means that the effect of democracy by itself is positive towards conflict initiation, but the effect of democracy becomes negative towards conflict with higher GDPPC. This is an interesting find considering that GDPPC had an unexpected positive effect on conflict in step 2. It would seem that it is the combination of democracy and GDPPC that explains conflict initiation best, while it makes no contribution to explaining the number of counts. As either the interaction or basic terms of the interaction are insignificant, the relationship between GDPPC and democracy does not explain the dual effect of democracy. It would seem that this effect is not due to analytical problems. The outlier term included in the model explaining military violence is insignificant in the logistic and significant in the count part of model 17. The effect of the outlier variable is very strong relative to the other variables with a coefficient of 3.46. Democracy is no longer significant in the model. However the insignificance of democracy is due to the interaction with GDPPC, when I tried to remove this interaction from the regression democracy was significant along with the outlier variable. This means the outliers do not explain the entire effect of democracy only a large part of it.

Table 8: Regression Coefficients from Step 3

	Protests		Violent Protests		Small-Scale Violence		Asymmetric Violence		Military Violence		Massive Unconventional Violence	
	Model 13		Model 14		Model 15		Model 16		Model 17		Model 18	
	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log
(Ln)Population	0.552***	-1.265***	0.514***	-1.368***	0.477***	-1.208***	0.551***	-1.535***	0.566***	-1.428***	0.111***	-0.675***
	(0.0168)	(0.101)	(0.0170)	(0.109)	(0.0203)	(0.0875)	(0.0259)	(0.146)	(0.0192)	(0.123)	(0.0424)	(0.0708)
(Ln)GDPpc	0.382***	-0.686***	0.191***	-0.832**	-0.0833**	-0.923***	0.147***	-1.056***	-0.0818**	-1.616***	-0.227**	0.353**
	(0.0423)	(0.157)	(0.0383)	(0.341)	(0.0394)	(0.220)	(0.0468)	(0.334)	(0.0330)	(0.385)	(0.115)	(0.149)
Anocracy	0.0901	2.071***	0.268***	-2.526**	0.566***	-1.478*	0.462***	-2.981**	0.647***	-0.197	-0.331	-1.230***
	(0.114)	(0.555)	(0.0975)	(1.080)	(0.117)	(0.839)	(0.137)	(1.377)	(0.112)	(0.877)	(0.296)	(0.405)
Democracy	-0.645	-5.019***	-0.550	-2.991	-0.306	-5.199***	0.0577	-5.571**	-0.0890	-8.752***	-0.537	4.688***
	(0.430)	(1.766)	(0.386)	(2.657)	(0.417)	(1.857)	(0.504)	(2.572)	(0.357)	(3.209)	(1.151)	(1.471)
Ethnicity	0.156***	0.144	0.241***	0.230	0.395***	0.102	0.448***	0.248	0.485***	0.198	0.163*	-0.208*
	(0.0329)	(0.148)	(0.0309)	(0.147)	(0.0376)	(0.123)	(0.0444)	(0.159)	(0.0345)	(0.177)	(0.0925)	(0.114)
State Instability	0.941***	-0.394	0.667***	-1.047**	0.666***	-0.738**	0.579***	-0.717	0.685***	-0.724	-0.816***	-1.602***
	(0.101)	(0.353)	(0.0925)	(0.439)	(0.107)	(0.339)	(0.134)	(0.497)	(0.0954)	(0.512)	(0.282)	(0.459)
Internet Bias	0.0945***	-0.124***	0.088***	-0.089***	0.065***	-0.077***	0.099***	-0.031**	0.062***	-0.140***	0.061***	-0.039***
	(0.00283)	(0.0126)	(0.00267)	(0.0133)	(0.00318)	(0.0105)	(0.0035)	(0.0134)	(0.00279)	(0.0190)	(0.0119)	(0.0140)
GDP*Dem	0.0890*	0.534**	0.0703	0.462	0.115**	0.747***	0.0736	0.780**	0.0919**	1.257***	0.0839	-0.614***
	(0.0504)	(0.209)	(0.0457)	(0.358)	(0.0490)	(0.242)	(0.0587)	(0.345)	(0.0418)	(0.434)	(0.140)	(0.178)
Instability*Anocracy	-0.519***	-37.71*	-0.195	1.083	-0.0197	0.692	0.618**	1.165	0.174	-1.054	1.730***	1.374*
	(0.200)	(21.37)	(0.190)	(1.658)	(0.223)	(1.185)	(0.259)	(1.911)	(0.209)	(1.615)	(0.536)	(0.767)
Ethnicity2	-0.00945**	-0.0205	-0.0155***	-0.00893	-0.0380***	-0.0138	-0.037***	-0.0117	-0.044***	-0.00469	-0.0129	0.0168
	(0.00464)	(0.0220)	(0.00432)	(0.0189)	(0.00510)	(0.0176)	(0.00662)	(0.0221)	(0.00485)	(0.0230)	(0.0119)	(0.0153)
Outliers					2.901***	-15.28			3.467***	-15.57		
					(0.505)	(8,365)			(0.531)	(27,128)		
Constant	-8.632***	17.14***	-6.204***	17.06***	-4.098***	17.26***	-7.394***	19.82***	-3.184***	22.60***	-0.546	6.320***
	(0.388)	(1.720)	(0.341)	(2.808)	(0.390)	(1.960)	(0.466)	(3.120)	(0.340)	(3.206)	(0.940)	(1.221)
Observations	4,697		4,697		4,697		4,697		4,697		4,697	4,697

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p

The interaction term between anocracy and the instability function is only an improvement in predicting counts of model 16, asymmetric violence. Instability has additional positive effect towards asymmetric events for anocracy. This means that the unexpected findings in regard to democracy cannot be attributed to anocracies.

The squared term of ethnicity is insignificant towards explain excess zeroes in all models, but significant in explaining counts in all models, except model 18 massive unconventional violence. Recalling the results of step 2 ethnicity had a dual effect in the models predicting violent protest, asymmetric and military violence. In these models we see that this dual effect can be explained by curvilinearity in these models. In all the models the basic term is positive while the interaction is negative. This means that these conflict types are more common in countries where the excluded part of the population is smaller, at a certain size the effect of ethnicity flattens.

Finally, I specified a whole range of models where I tested for an interaction between democracy and the internet bias. However, only a few if the models were significant, and those who were predicted that democracies gained more zero counts with time relative to other regime types. The unexpected findings regarding democracy and violent conflict can therefore not be explained by the internet bias. For the regression table see Table A4 in appendix D.

Now the only unexpected find from step 2 that remains is the unexpected difference in strength between democracy and anocracy for small-scale and military violence. In both step 3 models the outlier term is significant and strongly positive. The democracy term is also as mentioned insignificant, but not due to the outliers. In order to isolate the effect of the outliers I specified the models without the insignificant interactions with regime. The coefficients and incident rate ratios can be seen in table 8.

Table 9: Coefficients and Incident Rate Ratios for Respecified versions of model 15 and 17.

	Small-scale Violence Coefficients		Military violence Coefficients		Small Scale Violence Incidence Rate Ratios	Military Violence Incidence Rate Ratios
	Model 15 Modified		Model 17 Modified		Model 15 Modified	Model 17 Modified
	Nb	Log	Nb	Log	Irr	Irr
Democracy	0.657***	0.641***	0.665	0.814	1.897	1.946
	(0.212)	(0.064)	(.057)	(.313)	(0.122)	(0.112)
Anocracy	-0.528	0.624***	0.749	0.045	1.866	2.116
	(0.407)	(0.088)	(.082)	(.671)	(0.164)	(0.174)
GDPPC	-0.320***	0.003	-0.0164	-0.703	1.003	0.983
	(0.0819)	(0.025)	.(022)	(.145)	(0.025)	(0.021)
Population	-1.106***	0.471***	0.566	-1.315	1.602	1.762
	(0.0739)	(0.020)	(.019)	(.110)	(0.032)	(0.034)
Ethnicity	0.0720	0.391***	0.482	.053	1.478	1.619
	(0.123)	(0.037)	(.034)	(.170)	(0.055)	(0.056)
State Instability	-0.867***	0.603***	0.670	-1.594	1.827	1.955
	(0.314)	(0.092)	(0.084)	(.579)	(0.168)	(.165)
Internet Bias	-0.0824***	0.063***	0.060	-0.149	1.065	1.062
	(0.00974)	(0.003)	(0.002)	(.019)	(0.003)	(0.002)
Ethnicity2	-0.0155	-0.038***	-0.043	0.011	0.962	0.957
	(0.0185)	(0.005)	(0.004)	(0.022)	(0.004)	(0.004)
Outlier	-16.16	2.932***	3.506	-15.083	18.763	33.346
	(13.142)	(0.500)	(0.530)	(28735.5)	(9.380)	(17.705)
Constant	12.04***	-4.675***	-3.687	15.404	0.009	0.025
	(0.952)	(0.293)	(0.260)	(1.543)	(0.002)	(0.006)
N	4697		4697		4697	4697

Standard errors in parenthesis
*p<0.10, **p<0.05, ***p<0.01

With the outlier variable in the regression, the effect of democracy is now smaller then that of anocracy for military violence while it remains stronger for small-scale violence. It would seem then that though the outliers have a strong impact on the effect of democracy they cannot explain why democracy has a larger prediction of counts than anocracy in step 2. If the inclusion of the outlier variables had made the democracy variable insignificant I would be able to confidently claim support for *H2*. However, given the results, support is mixed. Part of the reason for this is that while these counts are very high and do influence the effect of democracy considerably one may still ask whether it is fair to consider them abnormal. After all, these observations are democracies and the do have high counts. Democracies still have more excess zero values, but the high count erodes support for *H2*.

5.0 Conclusion

I have conducted an analysis of regime and conflict using event data and zero inflated negative binomial regression models. The purpose of the paper was to disaggregate conflict and see if findings from previous research would persist. The research question of this thesis was whether there are any kinds of conflict that are particularly common in democracies. Based on the analysis done here I find that non-violent protest actions are more common in democracies than other regimes. I also find that democracies have more counts of small-scale violence and military violence than other regimes. At the same time, democracies have more excess zero values of these conflict types and the counts are heavily affected by influential outliers. I addition the difference between democracies and anocracies are in any case not very large. Therefore, I cannot claim that these conflict types are particularly common in democracies. Therefore, while I did not find the exact relationship between democracy, anocracy and conflict predicted from theory, neither do I find relationships that go directly contract to them.

The most persistent finding in the thesis is that authoritarian regimes have fewer conflicts than other regimes. They have both more excess zeroes and fewer counts for all conflicts.

Table 10: Summary of Findings

Hypothesis	Step 1	Step2	Step 3
H1	Supported	Supported	Supported
H2	Mixed Evidence	Mixed Evidence	Mixed Evidence
H3	Mixed Evidence	Mixed Evidence	Mixed Evidence
H4	Supported	Supported	Supported

In the introduction, I listed testing the feasibility of automatically coded event data in aiding the disaggregation of variables in conflict studies. It seems that it can indeed aid in disaggregating conflict, but there are several problems with this kind of data that needs to be accounted for. In the end it is a matter of how much one is willing to trust that the computer puts the events in the right category. In retrospect I should have taken greater steps to assure that this was the case, as it stands now I have no way of discovering with certainty weather strange patterns in the data are

natural or the result of coding bias.

Also while this paper goes to significant lengths to disaggregate conflict it does little to disaggregate actors, regimes or spatial relationships. Event data have the potential to aid in the disaggregation effort. Disaggregating conflicts by type as done in this paper together with disaggregation of spatial dimensions to a sub-national level, and a clearer distinction of actors and their objectives will probably reveal relationships that are invisible at the monadic level. Such aggregation has been done before (Buhaug et al 2009, Collier & Rhoner 2008), but event data could potentially hold the key to doing it all at once. On a related note it may be that analysis using event data of this nature would benefit from including data on press and/or internet coverage as a control variable, given the large increase of counts towards the later years of the data and the nature of news as a source.

Further research should expand upon the unexpected findings regarding democracy. Particularly events of the type included in the small-scale violence category. Event data allows the possibility of studding such low scale conflict together with non-violent political engagement and larger conflicts. With further disaggregation of conflicts, regimes and spatial dimensions, this could help explain in which situations dissidences choose which forms of political engagement whit greater accuracy than previous research.

The positive relationship between democracy and small-scale violence that I find is only a correlation in one dataset. The small-scale category also includes several different types of conflict, which could possibly be governed by different logics. Case studies of democracy and these conflict types, or more narrowly focused statistical studies could add to this study by discovering the causal mechanisms that govern these conflicts. One particular part of the regime variable that ought to be disaggregated is the difference between institutional and liberal democracy. Some claim (Mousseau 2009) that the democratic peace is caused be liberal norms and liberal trade rather than institutional characteristics of democracies. It could be that the relationship between regime and small-scale conflict is tied to particular regime forms rather than the overarching typology of democracy, anocracy and authoritarian regimes.

6.0 References

Ackerman, Peter & Karatnycky, Adrian (2005) "How Freedom is Won: From Civic Resistance to Durable Democracy" Downloaded 13.04.14 from: http://agnt.org/snv/resources/HowFreedomisWon.pdf

Aristotle [Unknown] (1996) *The Politics and The Constitution of Athens,* Revised Student Edition, Cambridge: Cambridge University Press.

Babst, Dean V. (1964). "Elective Governments - a Force for Peace", in The Wisconsin Sociologist, vol. 3, Iss. 1, pp. 9-14.

Bagozzi, Benjamin E. (2011) "Forecasting Civil Conflict with Zero-Inflated Count Models", available from: http://www.parusanalytics.com/eventdata/papers.dir/Bagozzi.ZICM.2011.pdf

Buhaug & Gates (2002) "The Geography of Civil War", in *Journal of Peace Research,* Vol. 39, No. 4, pp. 413-433.

Buhaug, Halvard (2006) "Relative Capability and Rebel Objective in Civil War" in *Journal of Peace Research* Vol. 43, Iss. 6, pp. 691-708.

Buhaug, Halvard (2010) "Dude, where's my conflict? LSG, relative strength, and the location of civil war.", in *Conflict Management and Peace Science* Vol.27, Iss. 2, pp. 107-128.

Buhaug, Halvard, & Rød, Jan K. (2006)"Local determinants of African civil wars, 1970–2001." *Political Geography,* Vol.25 Iss.3, pp. 315-335.

Buhaug, H., & Lujala, P. (2005) "Accounting for scale: Measuring geography in quantitative studies of civil war", in *Political Geography,* Vol.24, Iss.4, pp. 399-418.

Cole, Benjamin R. & Marshall, Monty, G. (2011) "Global Report 2011: Conflict, Governance, and

State Fragility", Retrieved 29.10.13 from: http://www.systemicpeace.org/GlobalReport2011.pdf

Collier, Paul & Hoeffler, Anke (1998) "On Economic Causes of Civil War", in *Oxford Economic Papers,* Vol. 50, Iss.4 pp. 563-573.

Collier, Paul, & Anke Hoeffler. (2004) "Greed and Grievance in Civil War." in, *Oxford economic papers,* Vol. 56. Iss.4 pp. 563-595.

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

Collier, Paul, & Dominic Rohner. (2008) "Democracy, development, and conflict", in *Journal of the European Economic Association* Vol.6 Iss.2-3 pp. 531-540.

Cunningham, Kathleen G. (2013) "Understanding Strategic Choice: The Determinants of Civil War and nonviolent Campaign in Self-Determination Disputes", in *journal of Peace Research,* Vol. 50, Iss. 3, pp291-304.

Cederman, Lars-Erik, Gleditsch, Kristian. S and Hug, Simon. (2013). "Elections and ethnic civil war", in *Comparative Political Studies,* Vol.46, Iss, 3, pp.387-417.

Cederman, Lars-Erik, Hug, Simon. and Krebs, Lutz. F. (2010a) "Democratization and civil war: Empirical evidence", in *Journal of Peace Research*, Vol.47, Iss.4, pp. 377-394.

Cederman, Lars-Erik, Wimmer, Andreas and Min, Brian. (2010b) "Why do ethnic groups rebel?: New data and analysis", *World Politics*, Vol.62, Iss.1, 87-119.

Cederman, Lars-Erik, Wimmer, Andreas and Min, Brian (2009) "Ethnic politics and armed conflict: a configurational analysis of a new global data set." in, *American Sociological Review,* Vol.74, Iss.2, pp. 316-337.

Danilovic, Vesna & Clare, Joe (2007) "The Kantian Liberal Peace (Revisited)", in *American Journal of Political Science,* Vol. 51, Iss 2, pp. 397- 414.

Davenport, Christian (2001) "Multidimensional Threat Perception and State Repression: An Inquiry Into why States Apply Negative Sanctions", I *American Journal of Political Science*, Vol.39, Iss. 3, pp. 683-713.

Davies, James C. (1962) "Toward a Theory of Revolution", in *American Sociological Review,* Vol. 27, Iss. 1, pp. 5-19.

Dixon, Jeffrey. (2009) "What causes civil wars? Integrating quantitative research findings." *International Studies Review*, Vol.11, Iss.4, pp. 707-735.

Doyle, Michael W. (1986) "Liberalism and World Politics" *The American Political Science Review",* Vol. 80, Iss. 4, pp. 1151-1169.

Eck, Kristine & Hultman, Lisa (2007) "One-Sided Violence Against Civilians in War: Insights from New Fatality Data", in *Journal of Peace Research,* Vol. 44, Iss. 2, pp. 233-246.

Ellingsen, Tanja (2000) "Colorful community or ethnic witches' brew? Multiethnicity and domestic conflict during and after the Cold War." in, *Journal of Conflict Resolution,* Vol.44, Iss.2, pp. 228-249.

Enia, Jason. Zeynep, Taydas and James, Patrick. (2011) "Why do civil wars occur? Another look at the theoretical dichotomy of opportunity versus grievance." in, *Review of International Studies.* Vol.37 Iss.05 pp. 2627-2650.

Fearon, James D. and Laitin, David D. (2003) "Ethnicity, insurgency, and civil war.", in *American political science review,* Vol.97, Iss.1, pp. 75-90.

Fjelde, Hanne (2010) "Generals, Dictators and Kings Authoritarian Regimes and Civil Conflict, 1973-2004", in *Conflict Management and Peace Science,* Vol. 27, No. 3, pp.195-218.

Freedom House (2013) "Freedom in The World", Available From: http://www.freedomhouse.org/sites/default/files/FIW%202013%20Booklet_0.pdf

Gartzke, Erik (2007) "The capitalist peace", in *American Journal of Political Science* Vol.51, Iss.1, pp. 166-191.

Gates, Scot. Moses, Johnathon and Knudsen, Torbjørn (1996) "Democracy and peace: A more Sceptical view", in *Journal of Peace Research,* Vol. 33, Iss. 1, pp. 1-10.

Gleditsch, Kristian S. (2002) "Expanded Trade and GDP Data", in *The Journal of Conflict Resolution,* Vol. 46, Iss. 5, pp. 712-724.

Gleditsch, Kristian S, & Ruggeri, Andrea (2010) "Political opportunity structures, democracy, and civil war.", in *Journal of Peace Research,* Vol.47, Iss.3, pp. 299-310.

Gleditsch, Kristian S. Metternich, Nils W & Ruggeri, Andrea (2014) "Data and progress in peace and conflict research", in *Journal of Peace research,* Vol. 51, Iss. 2, pp. 301-314.

Gleditsch, Nils P. & Hegre, Håvard (1997) "Peace and Democracy: Three Levels of Analysis", in *Journal of Conflict Resolution,* Vol. 41, Iss. 2, pp. 283-310.

Gurr, Ted, R. (1970) *Why Men Rebel,* Princeton: Princeton University Press.

Gurr, Ted, R. [1970] (2011) *Way Men Rebel,* Fortieth Anniversary Edition, Boulder: Paradigm Publishers.

Gurr, Ted R. (2000) *Peoples Versus States: Minorities at Risk in The New Century",* Washington D.C: Institute of Peace.

Hegre, Håvard & Sambanis, Nicholas (2006) "Sensitivity Analysis of Empirical Results on Civil War Onset", in *The Journal of Conflict Resolution,* Vol. 50, Iss. 4, pp. 508-535.

Hegre, Håvard, Ellingsen, Tanja, Gates, Scott and Gleditsch, Nils P. (2001) "Toward a Democratic Civil Peace? Democracy, Political Change and Civil War, 1816-1992", in American Political Science Review, Vol. 95, Iss. 1, pp. 33-48.

Henderson, Errol, A. & Regan, Patrick, M. (2013) "Democracy, Threats and Political Repression in Developing Countries: Are Democracies Internally Less Violent?", Retrieved 02.08.13, from: http://cdp.binghamton.edu/papers/threats-full.pdf

Hendrix, Cullen S. (2010) "Measuring State Capacity: Theoretical and empirical implications for the study of civil conflict", in *Journal of Peace Research,* Vol. 43, Iss. 3, pp. 273-285.

Hibbs, Dougglas A. (1973) *Mass Political Violence: A Cross-National Analysis,* New York: Wiley.

Hunziker, Philipp (2013) "GROW^{up} Research Front-End Documentation", Release 1.1, [Dataset and Documentation], Available From: http://www.icr.ethz.ch/data

Kant, Immanuel (1795) *Perpetual Peace: A Philosophical Essay*, retrieved 28.10.13 form: http://www.archive.org/stream/perpetualpeaceph00kantuoft/perpetualpeaceph00kantuoft_djvu.txt

King, Garry (2011) "Ensuring the Data-Rich Future of the Social Sciences", in *Science,* Vol.331, Iss. 6018 pp. 719-721.

Krain, Matthew & Merrisa, Edson Myers (1997) "Democracy and Civil War: A Note on The Democractic Peace Proposition", in *International Interactions,* Vol. 23, Iss. 1, pp. 109-118.

Krain, Matthew (1998) "Contemporary Democracies Revisited: Democracy, Political violence, and Event Count Models", in *Comparative Political Studies,* Vol. 31, Iss 2, pp.139-164.

Kruger, Lacek & Tammen, Ronald (2012) *Performance of Nations,* Lanham: Rowman and Littlefield.

Leetaru, Kalev and Schrodt, Philip. (2013). "GDELT: Global Data on Events, Language, and Tone", 1979-2012, International Studies Association Annual Conference, April 2013. San Diego, CA.

Li, Quan (2005) "Does Democracy Promote or Reduce Transnational Terrorist Incidents?" *Journal of Conflict Resolution,* Vol.49**,** Iss.2, 278-297.

Linz, Juan J. & Stepan, Alfred (1996) *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe,* London: The Johns Hopkins University Press.

Machiavelli, Niccolo ([1531] 1950) *The Prince & The Discourses on The First Ten Books of Titus Levy,* New York: Random House.

Maoz, Zeev & Russet, Bruce (1993) "Normative and Structural Causes of Democratic Peace" in *The American Political Science Review,* Vol. 87, Iss. 3, pp. 624-638.

Markus, Gregory B. and Nesvold, Betty A. (1972) "Governmental Coerciveness and Political Instability: An Exploratory Study of Cross-National Patterns", in *Comparative Political Studies,* Vol.5, Iss.2, pp.231-244.

Marshall, Monty. G., & Jaggers, K. (2013) "Polity IV Annual Time-Series 1800-2012" [Dataset & Codebook] Retrieved 21.10.13 from: http://www.systemicpeace.org/inscr/inscr.htm

McAddam, Doug McCarthy, John D. & Zald, Mayer (1996) "Introduction: Opportunities, mobilizing structures, and framing processes – towards a synthetic, comparative perspective on social movements", in McAddam, Doug McCarthy, John D. & Zald, Mayer, (red.) *Comparative Perspectives on Social Movements,* Cambridge: Cambridge University Press, pp. 1-20.

Minorities at Risk (2014) http://www.cidcm.umd.edu/mar/

Moore, Will H. (1998) "Repression and Dissent: Substitution, Context and Timing", in *American Journal Of Political Science,* Vol. 42, Iss. 3, pp, 851-873.

Moses, Jonathon & Knudsen, Torbjørn L. (2007) *Ways of Knowing* [Second Edition], Houndmills: Palgrave MacMillan.

Mousseau, Michael (2009) "The Social Market Roots of Democratic Peace", in *International Security,* Vol. 33, Iss. 4, pp. 52-86.

Muller, Edward N & Wedee, Erich (1990) "Cross-National Variation in Political Violence: A Rational Action Approach", in *The Journal of Conflict Resolution,* Vol. 34, Iss. 4, pp. 624-651.

Oberschall, Anthony (1996) "Opportunities and Framing in the Eastern European revolts of 1989" in McAddam, Doug McCarthy, John D. & Zald, Mayer, (red.) *Comparative Perspectives on Social Movements,* Cambridge: Cambridge University Press, pp. 94-121.

Piazza, James A. (2008) "Do Democracy and Free Markets Protect Us From Terrorism?", in *International Politics,* Vol. 45, pp. 72-91.

Plato [Approx. 308bc] (2001) *Platon: Samlede Verker: Bind 5: Kleitofon og Staten,* Oslo: Vidarforlaget.

Powell, Bingham G. Jr. (1982) *Contemporary Democracies,* Massachusetts: Harvard University Press.

Rummel, R. J (1995) "Democracy, Genocide and Mass Murder", in *Journal of Conflict Resolution,* Vol.39, Iss. 3, pp. 3-25.

Rummel, R. J (1997) *Power Kills: Democracy as a Method of Nonviolence,* New Brunswick: Transaction Publishers.

Sandler, Todd (2014) "The Analytical Study of Terrorism: Taking Stock", in *Journal of Peace Research,* Vol. 51, Iss. 2, pp. 257-271.

Schneider, Gerald (2014) "Peace through globalization and capitalism? Prospects of two liberal propositions", in *Journal of Peace Research,* Vol. 51, Iss.2, pp. 173-183.

Schneider, Gerald, and Gleditsch Nils P. (2014) "The capitalist peace: The origins and prospects of a liberal idea." in, *International Interactions,* Vol.36, Iss.2, pp. 107-114.

Schrodt, Philip A. (2014) "Seven deadly sins of contemporary quantitative political analysis", *Journal of Peace Research,* Vol. 51, Iss. 2, pp. 287-300.

Schrodt, Philip, A. (2011) "TABARI: Textual Analysis by Augmented Replacement Instructions", Version 0.7.6, available from: http://eventdata.parusanalytics.com/tabari.dir/tabari.manual.0.7.6.pdf

Schrodt, Philip, A. (2012) "CAMEO: Conflict and Mediation Event Observations Event and Actor Codebook" Version: 1.1b3, Available from: http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf

Sobek, David & Payne, Caroline (2010) "A Tale of Two Types: Rebel Goals and The Onset of Civil Wars", in *International Studies Quarterly,* Vol. 53, Iss. 1, pp 213-240.

Stephan, Maria J. & Chenoweth, Erica (2008) "Why Civil Resistance Works: The Strategic Logic of Nonviolent Struggle", in *International Security,* Vol. 33, Iss.1, pp. 7-44.

Strand, Håvard. Hegre, Håvard, Gates, Scott. And Jones, Mark P. (2006) "Institutional Inconsistency and Political Instability: Polity Duration, 1800–2000", in *American Journal of Political Science,* Vol. 50, Iss. 4, pp. 893.908.

Tarrow, Sidney (1996) "Social Movements and the state: Thoughts on the policing of protest", in McAddam, Doug McCarthy, John D. & Zald, Mayer, (red.) *Comparative Perspectives on Social Movements,* Cambridge: Cambridge University Press, pp. 62-92.

Taylor, Bob (2013) "Swedish Riots Could Represent Europe's Future", *The Washington Times,* viewed 25.11.13 at: http://communities.washingtontimes.com/neighborhood/what-world/2013/may/29/swedish-riots-could-represent-europes-future/

Theis, Cameron (2010) "Of rulers, rebels, and revenue: State capacity, civil war onset, and primary commodities", in *Journal of Peace Research,* Vol. 47, Iss. 3, pp 321-332.

Tilly, Charles (1978) *From mobilization to revolution,* New York: McGraw-Hill.

Tocqueville, Alexis de. ([1840] 2001) *Democracy in America,* New York, New American Library.

US Census Bureau "World Population Clock" Available from: http://www.census.gov/popclock/

Vanhanen, Tatu (2003) "Democratization and power resources, 1850-2000", [computer file] FSD1216, version 1.0 (2003-03-10). Tampere: Finnish Social Science Data Archive. http://www.fsd.uta.fi/en/data/catalogue/FSD1216/meF1216e.html

Vreeland, James Raymond (2008) "The effect of political regime on civil war unpacking anocracy", in *Journal of Conflict Resolution,* Vol.52, Iss.3, pp. 401-425.

Wagner, Daniel "What Sweden's Riots Say About Europe's Rising Social and Political Risks", *The World Post,* viewed 12.05.14 at: http://www.huffingtonpost.com/daniel-wagner/what-swedens-riots-say-ab_b_3334501.html

Walsh, John (2003) "Egypt's Muslim Brotherhood: Understanding Centrist Islam" In *Harvard International Review,* Vol. 24, Iss. 4, pp 32-36.

Østerud, Øyvind (2007) *Norsk Statsvitenskapelig Leksikon,* Oslo: Universitetsforlaget.

Appendices

Appendix A: List of Conflict Variables.

Variable Name	CAMEO eventcode	Description	Category
V3	140	Protest, political dissent, not specified	Protest
v4	141	Demonstrate not specified	Protest
v5	1411	Demonstrate for Leadership Change	Protest
v6	1412	Demons, policy Change Demonstrate for Policy Change	Protest
v7	1413	Demonstrate for rights	Protest
v8	1414	Demonstrate for Change in Regime	Protest
v9	1421	Hunger strike for leadership change	Protest
v10	1422	Hunger strike for policy change	Protest
v11	1424	Hunger strike for change in institutions/regime	Protest

v12	1431	Strike for leadership change	Protest
v13	1432	Strike for policy change	Protest
v14	1434	Strike for change in institutions/regime	Protest
v15	1441	Obstruct Passage for leadership change	Protest
v16	1442	Obstruct Passage for policy change	Protest
v17	1444	Obstruct Passage for change in institutions/regime	Protest
v18	1451	Protest Violently for leadership change	Violent Protest
v19	1452	Protest Violently for policy change	Violent Protest
v20	1454	Protest Violently for rights	Violent Protest
v21	171	Size or damage property	Violent Protest
v22	1711	Confiscate property	Violent Protest
v23	1712	Destroy property	Violent Protest
v24	173	Arrest or Detain	Violent Protest
v25	176	Attack Cybernetically	Excluded
v26	180	Use unconventional violence, not specified	Small-Scale Violence
v27	181	Abduct, Hijack or take hostage	Asymmetric Violence
v28	182	Physically assault, not specified	Small-Scale Violence
v29	1821	Sexually Assault	Small-Scale Violence
v30	1822	Torture	Small-Scale Violence
v31	1823	Kill, without weaponry	Small-Scale Violence
v32	183	Non- Military bombing, not specified	Asymmetric Violence
v33	1831	Suicide Bombing	Asymmetric Violence
v34	1832	Vehicular bombing	Asymmetric Violence
v35	1833	Roadside Bombing	Asymmetric Violence
v36	1834	Location Bombing	Asymmetric Violence
v37	184	Use Human Shields	Asymmetric Violence
v38	185	Attempted Assassination	Asymmetric Violence
v39	186	Successful Assassination	Asymmetric Violence
v40	190	Conventional Military force, not specified	Military Violence
v41	191	Impose Blockade	Military Violence
v42	192	Occupy territory	Military Violence
v43	193	Fight with small arms, and light weapons	Military Violence
v44	194	Fight with artillery and armor	Military Violence
v45	195	Employ aerial Weapons, not specified	Military Violence
v46	1951	Precision-guided munitions	Military Violence
v47	1952	UAVs	Military Violence
v48	196	Violate Ceasefire	Excluded
v49	200	Massive Unconventional Force	Mass Violence
v50	201	Mass Expulsion	Mass Violence

v51	202	Mass Killings	Mass Violence
v52	203	Ethnic Cleansing	Mass Violence
v53	204	WMDs, Not specified	Mass Violence
v54	2041	Use Chemical, biological or radiological weapons	Mass Violence
v55	2042	Nuclear Weapons	Excluded

Appendix B: Python Scripts

To export variables

```
import glob
import os


headerFile = open("headers.txt","r")
header = headerFile.readline().split("\t")

os.chdir("[unpacked GDELT directory]")

def fratil(a,b):
    ran = range(a,b+1)
    for i in range(0,len(ran)):
        ran[i] = str(ran[i])
        while len(ran[i]) < 3:
            ran[i] = "0"+ran[i]
    return ran


newFile = open("F****JoinedMk2.txt","w")


for yearFile in glob.glob("*.csv"):
    f = open("[unpacked GDELT directory ]"+str(yearFile),"r")
    for line in f:
        variables = line.split("\t")
        if variables[header.index("EventCode")] in []:
            if variables[header.index("Actor1Type1Code")] in
["INS","OPP","REB","SEP"]:
                newFile.write(line)
            elif variables[header.index("Actor1Type1Code")] in
["INS","OPP","REB","SEP"+ resten]:
                newFile.write(line)

    f.close()
    print "file complete: "+str(yearFile)



newFile.close()
print "DONE!"
```

To create the country year dataset

```
Created on Dec 17, 2013

@author: Jo
'''

# Create the empty array named "currentYear" (country as row, eventcodes as columns)
currentYear = {}
eventcodes =
["140","141","1411","1412","1413","1414","1421","1422","1424","1431","1432","1434","1
441","1442","1444","1451","1452","1454","171","1711","1712","173","176","180","181","
182","1821","1822","1823","183","1831","1832","1833","1834","184","185","186","190","
191","192","193","194","195","1951","1952","196","200","201","202","203","204","2041"
,"2042"]
eventDict = {}
for event in eventcodes:
    eventDict[event] = 0
countryFile = open("codes","r")
countryList = []
for code in countryFile:
    countryList.append(code[0:2])
    currentYear[code[0:2]] = dict(eventDict)


# Open the datafile, and create a new file called landaar.txt
dataFile = open("F:/GDELT_HISTORICAL/UNPACED/csv/Joined.txt","r")
newFile = open("F:/GDELT_HISTORICAL/UNPACED/csv/landaar2.txt","w")

endOfFile = False

# Iterate through all years
for year in range(1979,2014):
    print "Currently processing year: "+str(year)

    #Go through rows in the dataFile, until a row from the "next" year comes along
    while True:
        row = dataFile.readline()
        if not len(row):
            endOfFile = True
            variables = [0]*100
        else:
            variables = row.split("\t") #Split the raw text row into a list
containing all the variables

        if int(variables[3]) == year: # Check which year
            # Add 1 to the appropriate entry in currentYear
            country = variables[51] #ActionGeo_CountryCode
            eventcode = variables[26]
            try:
                currentYear[country][eventcode] += 1
                #print "ADDED: "+country+ "  "+eventcode
            except:
                pass
```

```
        else:
            # This row is the first of the new year, let's save the the previous
years data, and clear the "currentYear" array
            print "New year! Saving data to file..."

            #Saving this years data into newFile and clear it (in the same loop!!!)
            for country in countryList:
                eventdata = []
                for eventcode in eventcodes:
                    eventdata.append(str(currentYear[country][eventcode])) # List of
this country's eventdata (ex.  [0  2 3   56   8 4..] )
                    currentYear[country][eventcode] = 0 #Clear the entry...
                thisLine = country + "\t" + str(year) +"\t"+ "\t".join(eventdata) +
"\n" #Construct the row...
                newFile.write(thisLine)

            # Add 1 to the appropriate entry in currentYear (This is the next
year..!)
            country = variables[51]
            eventcode = variables[26]
            try:
                currentYear[country][eventcode] += 1
            except:
                pass

            #Go the next year!!
            break
    if endOfFile:
        newFile.close()
        dataFile.close()
        break
```

Header text, country codes and country years

Country codes

```
AA
AC
AE
AF
AG
AJ
AL
AM
AN
AO
AQ
AR
AS
AU
AV
AY
BA
BB
BC
```

BD
BE
BF
BG
BH
BK
BL
BM
BN
BO
BP
BR
BT
BU
BX
BY
CA
CB
CD
CE
CF
CG
CH
CI
CJ
CM
CN
CO
CQ
CS
CT
CU
CV
CW
CY
DA
DJ
DO
DR
EC
EG
EI
EK
EN
ER
ES
ET
EZ
FG
FI
FJ
FM
FO
FP
FR

GA
GB
GG
GH
GI
GJ
GK
GL
GM
GP
GQ
GR
GT
GV
GY
GZ
HA
HK
HO
HR
HU
IC
ID
IM
IN
IR
IS
IT
IV
IZ
JA
JE
JM
JO
KE
KG
KN
KR
KS
KT
KU
KV
KZ
LA
LE
LG
LH
LI
LO
LS
LT
LU
LY
MA
MB

MC
MD
MF
MG
MH
MI
MJ
MK
ML
MN
MO
MP
MR
MT
MU
MV
MX
MY
MZ
NC
NE
NF
NG
NH
NI
NL
NO
NP
NR
NS
NT
NU
NZ
OC
OD
OS
PA
PC
PE
PF
PG
PK
PL
PM
PO
PP
PS
PU
QA
RB
RE
RI
RM
RO
RP

RQ
RS
RW
SA
SB
SC
SE
SF
SG
SH
SI
SL
SM
SN
SO
SP
ST
SU
SW
SY
SZ
TD
TH
TI
TK
TN
TO
TP
TS
TT
TU
TV
TW
TX
TZ
UG
UK
UP
US
UV
UY
UZ
VC
VE
VI
VM
VQ
VT
WA
WE
WI
WS
WZ
YI
YM

```
ZA
ZI
```

Country code file

```
'''
Created on Mar 23, 2014

@author: Jo
'''

s = """GLOBALEVENTID    SQLDATE    MonthYear    Year    FractionDate    Actor1Code
Actor1Name    Actor1CountryCode    Actor1KnownGroupCode    Actor1EthnicCode
Actor1Religion1Code    Actor1Religion2Code    Actor1Type1Code    Actor1Type2Code
Actor1Type3Code    Actor2Code    Actor2Name    Actor2CountryCode
Actor2KnownGroupCode    Actor2EthnicCode    Actor2Religion1Code
Actor2Religion2Code    Actor2Type1Code    Actor2Type2Code    Actor2Type3Code
IsRootEvent    EventCode    EventBaseCode    EventRootCode    QuadClass
GoldsteinScale    NumMentions    NumSources    NumArticles    AvgTone
Actor1Geo_Type    Actor1Geo_FullName    Actor1Geo_CountryCode    Actor1Geo_ADM1Code
Actor1Geo_Lat    Actor1Geo_Long    Actor1Geo_FeatureID    Actor2Geo_Type
Actor2Geo_FullName    Actor2Geo_CountryCode    Actor2Geo_ADM1Code    Actor2Geo_Lat
Actor2Geo_Long    Actor2Geo_FeatureID    ActionGeo_Type    ActionGeo_FullName
ActionGeo_CountryCode    ActionGeo_ADM1Code    ActionGeo_Lat    ActionGeo_Long
ActionGeo_FeatureID    DATEADDED"""

list_s = s.split("    ")


for i in range(0,len(list_s)):
    if list_s[i] == "ActionGeo_CountryCode":
        print i
        break
```

Header text

```
GLOBALEVENTIDSQLDATE      MonthYear    Year   FractionDate Actor1Code   Actor1Name
      Actor1CountryCode   Actor1KnownGroupCode      Actor1EthnicCode
      Actor1Religion1Code Actor1Religion2Code Actor1Type1Code     Actor1Type2Code
      Actor1Type3Code     Actor2Code   Actor2Name   Actor2CountryCode
      Actor2KnownGroupCode      Actor2EthnicCode    Actor2Religion1Code
      Actor2Religion2Code Actor2Type1Code     Actor2Type2Code     Actor2Type3Code
      IsRootEvent  EventCode    EventBaseCodeEventRootCodeQuadClass
      GoldsteinScale      NumMentions  NumSources   NumArticles  AvgTone
      Actor1Geo_Type      Actor1Geo_FullName  Actor1Geo_CountryCode
      Actor1Geo_ADM1Code  Actor1Geo_LatActor1Geo_Long      Actor1Geo_FeatureID
      Actor2Geo_Type      Actor2Geo_FullName  Actor2Geo_CountryCode
      Actor2Geo_ADM1Code  Actor2Geo_LatActor2Geo_Long      Actor2Geo_FeatureID
      ActionGeo_Type      ActionGeo_FullName  ActionGeo_CountryCode
```

```
ActionGeo_ADM1Code  ActionGeo_LatActionGeo_Long     ActionGeo_FeatureID
DATEADDED
```

Appendix C: List of countries in Step 2

Country Code	Country Name
2	USA
20	Canada
40	Cuba
41	Haiti
42	Dominican Republic
51	Jamaica
52	Trinidad
70	Mexico
90	Guatemala
91	Honduras
92	El Salvador
93	Nicaragua
94	Costa Rica
95	Panama
100	Colombia
101	Venezuela
110	Guyana
115	Suriname
130	Ecuador
135	Peru
140	Brazil
145	Bolivia
150	Paraguay
155	Chile
160	Argentina
165	Uruguay
200	UK

205	Ireland
210	Netherlands
211	Belgium
212	Luxembourg
220	France
225	Switzerland
230	Spain
235	Portugal
255	Germany *Bundesrepublik only
290	Poland
305	Austria
310	Hungary
316	Czech Republic
317	Slovak Republic
325	Italia
339	Albania
340	Serbia
341	Montenegro
343	Macedonia
344	Croatia
345	Yugoslavia
347	Kosovo
349	Slovenia
350	Greece
352	Cyprus
355	Bulgaria
359	Moldova
360	Romania
365	Russia
366	Estonia
367	Latvia
368	Lithuania
369	Ukraine
370	Belarus
371	Armenia
372	Georgia
373	Azerbaijan
375	Finland
380	Sweden
385	Norway
390	Denmark
402	Cape Verde
404	Guinea-Bissau

411	Equatorial-Guinea
420	Gambia
432	Mali
433	Senegal
434	Benin
435	Mauritania
436	Niger
437	Ivory coast
438	Guinea
439	Burkina Faso
450	Liberia
451	Sierra Leone
452	Ghana
461	Togo
471	Cameroon
475	Nigeria
481	Gabon
482	Central African Republic
483	Chad
484	Congo Brazzaville
490	Congo Kinshasa
500	Uganda
501	Kenya
510	Tanzania
516	Burundi
517	Rwanda
520	Somalia
522	Djibouti
530	Ethiopia
531	Eritrea
540	Angola
541	Mozambique
551	Zambia
552	Zimbabwe
553	Malawi
560	South Africa
565	Namibia
570	Lesotho
571	Botswana
572	Swaziland
580	Madagascar
581	Comoros
590	Mauritius

600	Morocco
615	Algeria
616	Tunisia
620	Libya
625	Sudan
626	South Sudan
630	Iran
640	Turkey
645	Iraq
651	Egypt
652	Syria
660	Lebanon
663	Jordan
666	Israel
670	Saudi Arabia
679	Yemen *Republic of only
690	Kuwait
692	Bahrain
694	Qatar
696	UAE
698	Oman
700	Afghanistan
701	Turkmenistan
702	Tajikistan
703	Kyrgyzstan
704	Uzbekistan
705	Kazakhstan
710	China
712	Mongolia
713	Taiwan
731	Democratic Peoples Republic of Korea
732	Republic of Korea
740	Japan
750	India
760	Bhutan
770	Pakistan
771	Bangladesh
775	Myanmar
780	Siri Lanka
790	Nepal
800	Thailand
811	Cambodia
812	Laos

816	Vietnam
820	Malaysia
830	Singapore
840	Philippines
850	Indonesia
860	East Timor
900	Australia
910	Papua New Guinea
920	New Zealand
940	Solomon Islands
950	Fiji

Appendix D: Model Fit Tables

Table A 1 Tests for Significance of Models in Step 1.

Dependent variable	Alpha	Vuong Score	Pr>Z	Pr>chi-bar2	Prob>Chi^2
Protests (Model 1)	4.80	3.98	0.000	0.000	0.000
Violent Protests (Model 2)	3.88	2.39	0.000	0.000	0.000
Small Scale Violence (Model 3)	4.83	n/a	n/a	0.000	0.007
Asymmetric Violence (Model 4)	6.59	1.72	0.042	0.000	0.000
Military Violence (Model 5)	1.38	28.35	0.000	0.000	0.000
Mass Unconventional Violence (Model6)	9.75	2.07	0.019	0.000	0.088

Alpha = the log of the dispersion parameter.

Vuong Score and Pr>Z calculated from the Vuong test indicate the likelihood that a regular negative binomial model would be better suited than a zero inflated model.
Pr>chibar2 indicates the likelihood that a Poisson model would be better suited than a negative binomial model.
Chi^2 indicates the likelihood that the model is not statistically significant.

Table A2 Tests for Significance of Models in Step 2.

Dependent variable	Alpha	Vuong Score	Pr>Z	Pr>chi-bar2	Prob>Chi^2
Protests (Model 7)	1.62	8.46	0.000	0.000	0.000
Violent Protests (Model 8)	1.61	8.34	0.000	0.000	0.000
Small Scale Violence (Model 9)	2.08	8.86	0.000	0.000	0.000
Asymmetric Violence (Model 10)	2.94	7.81	0.000	0.000	0.000
Military Violence (Model 11)	2.35	6.35	0.000	0.000	0.000
Mass Unconventional Violence (Model12)	2.21	6.04	0.000	0.000	0.000

Alpha = the log of the dispersion parameter.
Vuong Score and Pr>Z calculated from the Vuong test indicate the likelihood that a regular negative binomial model would be better suited than a zero inflated model.
Pr>chibar2 indicates the likelihood that a zero inflated Poisson model would be better suited than a zero inflated negative binomial model.
Chi^2 indicates the likelihood that the model is not statistically significant.

Table A3 Tests for Significance of Models in Step 3.

Dependent variable	Alpha	Vuong Score	Pr>Z	Pr>chi-bar2	Prob>Chi^2
Protests (Model 13)	1.67	8.97	0.000	0.000	0.000
Violent Protests (Model 14)	1.61	8.44	0.000	0.000	0.000
Small Scale Violence (Model 15)	2.10	9.16	0.000	0.000	0.000
Asymmetric Violence (Model 16)	2.98	7.76	0.000	0.000	0.000
Military Violence (Model 17)	2.23	6.86	0.000	0.000	0.000
Mass Unconventional Violence (Model 18)	1.98	6.27	0.000	0.000	0.000

Alpha = the log of the dispersion parameter.
Vuong Score and Pr>Z calculated from the Vuong test indicate the likelihood that a regular negative binomial model would be better suited than a zero inflated model.
Pr>chibar2 indicates the likelihood that a zero inflated Poisson model would be better suited than a zero inflated

negative binomial model.
Chi^2 indicates the likelihood that the model is not statistically significant.

Appendix E Alternate Models

Table A4 Regression Models with an Interaction Term for the Internet Bias and Democracy

	Protests		Violent Protest		Small-Scale Violence		Asymmetric Violence		Military Violence		Unconventional Mass Violence	
	Model 19		Model 20		Model 22		Model 23		Model 24		Model 25	
	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log	Nb	Log
Population	0.548***	-1.069***	0.516***	-1.358***	0.512***	-1.097***	0.540***	-1.346***	0.565***	-1.341***	0.112***	-0.711***
	(0.0170)	(0.0762)	(0.0164)	(0.108)	(0.0203)	(0.0767)	(0.0259)	(0.110)	(0.0192)	(0.113)	(0.0430)	(0.0861)
GDPPC	0.428***	-0.339***	0.238***	-0.415***	0.0137	-0.308***	0.223***	-0.335***	-0.0155	-0.652***	-0.204***	-0.0967
	(0.0237)	(0.0822)	(0.0210)	(0.101)	(0.0258)	(0.0865)	(0.0305)	(0.0974)	(0.0220)	(0.134)	(0.0747)	(0.0947)
Anocracy	-0.00812	0.179	0.168**	-1.436	0.604***	0.0176	0.677***	-1.183*	0.725***	0.687	0.264	-1.137***
	(0.0883)	(0.373)	(0.0746)	(0.958)	(0.0917)	(0.536)	(0.113)	(0.701)	(0.0845)	(0.698)	(0.260)	(0.386)
Democracy	0.227*	-1.426***	0.491***	-0.233	0.865***	-0.277	0.905***	-0.135	0.819***	-0.0913	0.480	-0.237
	(0.126)	(0.377)	(0.102)	(0.453)	(0.121)	(0.375)	(0.143)	(0.491)	(0.101)	(0.433)	(0.498)	(0.674)
Ethnicity	0.149***	-0.0159	0.247***	0.238*	0.343***	-0.00391	0.444***	0.189	0.480***	0.00974	0.159*	-0.195*
	(0.0333)	(0.128)	(0.0307)	(0.141)	(0.0388)	(0.126)	(0.0450)	(0.152)	(0.0346)	(0.165)	(0.0962)	(0.118)
State Instability	0.784***	-1.159***	0.604***	-1.256***	0.618***	-0.971***	0.723***	-0.546	0.674***	-1.651***	-0.363	-1.124***
	(0.0877)	(0.378)	(0.0796)	(0.474)	(0.0945)	(0.365)	(0.115)	(0.447)	(0.0845)	(0.588)	(0.249)	(0.396)
Internet Bias	0.0965***	-0.161***	0.102***	-0.171***	0.0703***	-0.156***	0.106***	-0.0686***	0.0656***	-0.226***	0.0709***	-0.0277
	(0.00468)	(0.0196)	(0.00371)	(0.0348)	(0.00441)	(0.0274)	(0.00523)	(0.0225)	(0.00385)	(0.0367)	(0.0227)	(0.0289)
Internet bias*democracy	-0.00574	0.0762***	-0.0248***	0.102***	-0.00838	0.0967***	-0.0144**	0.0441	-0.00871*	0.120***	-0.0133	-0.00404
	(0.00570)	(0.0240)	(0.00474)	(0.0386)	(0.00561)	(0.0302)	(0.00665)	(0.0269)	(0.00497)	(0.0411)	(0.0227)	(0.0306)
Outliers									3.548***	-15.73		
									(0.531)	(25,453)		
Ethnicity2	-0.00830*	0.00257	-0.0165***	-0.00772	-0.0319***	3.64e-05	-0.0373***	-0.0119	-0.0433***	0.0219	-0.00878	0.0156
	(0.00473)	(0.0181)	(0.00436)	(0.0194)	(0.00535)	(0.0185)	(0.00671)	(0.0222)	(0.00489)	(0.0224)	(0.0123)	(0.0156)
Constant	-8.956***	13.35***	-6.825***	14.31***	-5.252***	12.38***	-8.008***	13.41***	-3.754***	15.60***	-1.112	9.765***
	(0.281)	(1.034)	(0.247)	(1.303)	(0.305)	(1.020)	(0.372)	(1.315)	(0.264)	(1.511)	(0.720)	(0.971)
Observations	4,697		4,697		4,697		4,697		4,697		4,697	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix F Download link to replication Data

http://www.mediafire.com/download/kkih7qjkdr7srbb/ReplicationJHN.rar