Bloodshed Technology? The Impact of Social Media Access on the Onset of Intrastate Conflict

Baier, Clemens

MASTER THESIS BY CLEMENS CARL BAIER
(S2158361)

BLOODSHED TECHNOLOGY? THE IMPACT OF SOCIAL MEDIA ACCESS ON THE ONSET OF INTRASTATE CONFLICT

Instructor: Dr. J. Masullo Jimenez

Second Reader: Dr. R. van der Haer

Program: MSc Political Science | International Politics

Wordcount: 9986

Universiteit Leiden
The Netherlands

SUBMISSION DATE: 13-06-2022

Abstract

While there is a growing body of research on the effects of information and communication technologies (ICTs) on the onset of intrastate conflict, the effects of social media access have not been analyzed thoroughly based on a large N, cross-country study. In this study, the impact of social media access on the onset of intrastate conflict is assessed, using country-year data (2000-2020) for 173 countries. It was hypothesized that social media access leads to an increase in the onset of intrastate conflict, ceteris paribus. Additionally, it was also hypothesized that the effect of social media access on the onset of intrastate conflict is stronger in more ethnically homogenous countries. These hypotheses were tested using a binomial logistic regression model. The main results of the analysis indicate there is only a significant correlation between Internet access and the onset of non-ethnic intrastate conflict. These results suggest that intrastate conflict is too broad of a category and should be disaggregated into smaller subgroups to develop more accurate empirical analyses. This research advances not only the study of intrastate conflict onset but also the academic discussion on the effects of social media access.

Table of Contents

1.Introduction ... 3
2.Literature Review ... 5
3.Theoretical Framework ... 7
4. Research Design ... 12
4.1 Case Selection and Models ... 12
4.2 Variables ... 13
4.2.1 Dependent Variable ... 13
4.2.2 Independent Variable ... 16
4.2.3 Control Variables ... 19
5. Analysis ... 21
5.1 Results ... 21
5.2.1 Hypothesis 1 ... 26
5. 2. 2 Hypothesis 2 ... 30
6. Conclusion and Future Research ... 34
7. References ... 36
8. Appendix ... 42

1. Introduction

The Internet and social media have drastically changed the way we live. Over 4.5 billion people use social media to connect with friends, share and gather information, and create their own content (Munger, 2019). This is not only a Western phenomenon. In Myanmar, for example, Facebook, the biggest and most widely used social media platform, has become synonymous with the Internet for a majority of people due to its low costs and the fact that it comes pre-installed on many cheap phones (Fink, 2018).

Consequently, the effects of social media have garnered the attention of social scientists, including conflict scholars, demonstrated by an ever-growing body of scientific research on the topic, especially in recent years. Additionally, headlines in newspapers such as “can social media change the course of war?” demonstrate the large interest in the effects of social media access on conflict and the importance that not only social scientists, but also the media, attribute to it (Tunzelmann, 2022). Academically, attempts have been undertaken to estimate the effects of social media and other forms of information and communication technology (ICT) on conflict (Pierskalla & Hollenbach, 2013; Tähtinen, 2021; Weidmann, 2015). However, these attempts were not based on large N cross-country analyses and lack generalizability. Therefore, this research will employ a large cross-country time series analysis to address whether social media access influences the onset of intrastate conflict.[1]

This thesis draws on the literature of intrastate conflict onset and the literature on the effects of social media, combining them in one theoretical framework. Based on this theoretical framework, two hypotheses are proposed: first, an increase in access to social media increases the likelihood of conflict onset, and second, this effect is expected to be stronger in countries

[1] For this research, intrastate armed conflict is defined as a “conflict between a government and a non-governmental party” that resulted in at least 25 battle related deaths within one calendar year (Gleditsch et al., 2002, p. 618).

that are more ethnically homogenous. Social media is expected to increase the likelihood for conflict onset due to its facilitation of collective action and its enhancement of grievances. Additionally, since within countries that are ethnically more homogenous, it is more likely that information is shared and received within an echo chamber, the anticipated effect of social media access on the onset of conflict is stronger due to group polarization taking place in such echo chambers. These two hypotheses will be tested using a binomial logistic regression model, using existent data from the World Bank and the GROW^{UP} project, as well as a unique dataset on social media penetration specifically compiled for this research (Vogt et al., 2021; World Bank, 2022a). The main finding, I argue, is that social media access does increase the likelihood for the onset of intrastate conflict, but only for conflict that is non-ethnic in nature.

Testing these two hypotheses adds to the growing body of literature on the effects of social media access, as well as contributes to the literature on the onset of intrastate conflict. Particularly, conducting a systematically large-N analysis will produce more generalizable results than previously conducted studies on the effects of social media, especially taking into account the unique large dataset compiled for this research. Additionally, due to the widespread usage of social media platforms nowadays, this topic is very relevant to society, not only due to the dangerous implications social media use could have on the onset of intrastate conflict, but also in order to determine the responsibility of such social media websites and policy makers in limiting the ability of users to spread dangerous content.

This research is divided into five sections. After a short literature review that provides an overview of the relevant literature on conflict onset, the theoretical framework of this research will be discussed, and the two hypotheses will be motivated. Afterwards, the empirical strategy is outlined, discussing both methods of analysis and data. Thereafter, an in-depth discussion of the main results is presented, and their implications are discussed. This thesis ends with a short conclusion, summarizing the findings and discussing possibilities for further research.

2. Literature Review

In the existing literature addressing the onset of intrastate conflict, two overarching strains can be identified. Early research focuses on the idea that groups within a state have grievances against another group, which then lead to the onset of intrastate conflict (Gagnon Jr, 1995; Gurr, 1993; Montalvo & Reynal-Querol, 2005; Sambanis, 2001). The general idea behind the grievance theory is that if a group is actively discriminated against, access to state power is not guaranteed, or the group is economically worse off; grievances against other groups occur (Denny & Walter, 2014). According to the literature, these grievances are necessary conditions for the onset of intrastate conflict.

Later research, especially starting with Collier and Hoeffler (2000) and Fearon and Laitin (2000), shifted the focus from grievances towards opportunities to rebel. Rooted in rational choice theory, these researchers analyzed whether it is beneficial and possible for a certain group to rebel or not. If doing so is deemed beneficial, and there are opportunities to rebel, the onset of intrastate conflict will become likely (Collier, Hoeffler, & Rohner, 2008). This literature argued that grievances alone will not lead to conflict. In the words of Collier and Hoeffler (2004 p. 588), "objective indicators for grievances add little explanatory power," while "a model that focuses on the opportunities for rebellion performs well." Following this lead, different factors related to economic opportunities have been identified as making the onset of intrastate conflict more likely. Such factors include the growth of the gross domestic product (GDP), the availability of natural resources the percentage of mountainous terrain, the size of the population that is male and between 15 and 29 years old, lack of access to education, and short-term water abundance, among others (Collier et al., 2008; Fearon & Laitin, 2003; Humphreys & Weinstein, 2008; Salehyan & Hendrix, 2014).

With a growing emergence of information and communication technologies (ICTs), such as mobile phones and the Internet, scholars have increasingly become interested in the effects of

ICTs on the onset of intrastate conflict (Adhami, 2007; Bailard, 2015; Howard & Hussain, 2011; Pierskalla & Hollenbach, 2013). The research on the effects of ICTs focuses both on the influence this new technology has on growing grievances, as well as on the increase in opportunities to rebel.

The majority of the studies found that access to ICTs increases the risk for the onset of intrastate conflict, due to an increase in the opportunities to rebel (Bailard, 2015; Diamond, 2010; Pierskalla & Hollenbach, 2013; Weidmann, 2015). The opportunities to rebel are increased because ICTs facilitate collective action by reducing communication problems, increasing the possibility to coordinate and to punish free riders.

For example, focusing on Africa, Pierskalla and Hollenbach (2013) and Bailard (2015) argued that cell phone coverage significantly increases the likelihood of conflict onset, because cell phones allow for the coordination of rebel groups and the overcoming of the free rider problem. Intrastate conflict, by definition a conflict against the government, often implies the fight against a stronger, better trained, and better-equipped enemy, which in turn makes fighting very costly (Bailard, 2015). However, overturning the government can benefit a larger population, which in turn makes free riding a central obstacle to collective action (Olson, 1989). Cell phones, due to almost instantaneous communication capabilities, allow for the control and punishment of free riders, and simultaneously reward participation (Baillard, 2015).

Additionally, ICTs allow rebel groups to learn from successful rebel groups in other states (Weidmann, 2015). Weidmann (2015) demonstrated that ICTs increase the opportunities for insurgencies by allowing rebel groups to learn from each other through the transnational spread of information, such as successful anti-government strategies. However, these studies, which emphasize the importance of opportunities, focus mainly on traditional forms of ICTs, such as (mobile) phones, or the Internet in general, but do not underscore the importance of social media specifically.

While much more literature focuses on the opportunities aspect of ICTs, some scholars have also argued that ICTs can fuel grievances (Adhami, 2007; Howard & Hussain, 2011, Weidmann 2015). Weidmann (2015), for example, argued that telecommunication networks can facilitate the spread of grievances between and within ethnic groups, which in turn increases the grievances in that group. Others have examined how the Internet can overcome government propaganda and highlight shared grievances between groups and group members (Adhami, 2007; Howard & Hussain, 2011). The Internet was used to build networks, which in turn were then used to spread information highlighting grievances across these networks. Adhami (2007) examined specifically the effects that the Internet can have on the recruitment of *jihadi* fighters by highlighting shared grievances, but as Weidmann (2015) demonstrated, grievances can spread through networks, regardless of the religion of the group.

While there is a lot of insightful literature about the effect of ICTs on the onset of intrastate conflict, there are three shortcomings in the literature. First, most of the literature is focused on a small number of cases or are single case studies. Second, while there is growing research on the effects of ICTs on the onset of intrastate conflict, most work is focused on (mobile) phones and the Internet in general, while a more disaggregated analysis of social media often does not take place. Lastly, most research is either focused on the opportunities to rebel or the increase in grievances. To overcome these limitations, this research employed a model with a large number of cases over a large number of years, focusing on social media specifically in order to identify the influence of social media on the onset of intrastate conflict, while considering the effects that social media potentially has on both the opportunities and the grievances.

3. Theoretical Framework

Information and communication technologies (ICTs) can make the onset of intrastate conflict more likely by facilitating and accelerating communication. Due to an increase in communication, rebel groups are able to learn from other rebel groups, overcome collective action and coordination problems, undermine government propaganda, and highlight shared grievances. As such, ICTs can influence both the opportunities to rebel and shared grievances (Adhami, 2007; Bailard, 2015; Howard & Hussain, 2011; Pierskalla & Hollenbach, 2013; Weidmann, 2015).

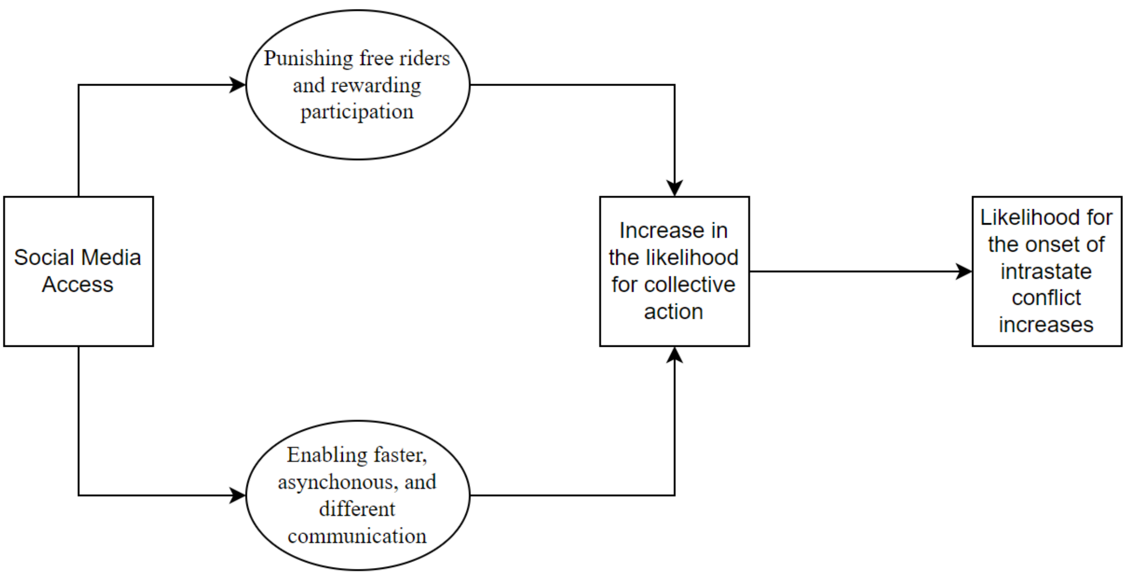
Social media platforms are currently the single largest platform of communication in the world (Munger, 2019). Social media platforms are not only facilitators of communication, but they also increase the speed and the scope of the information communicated, as well as change the type of the information being shared (Adhami, 2007). Additionally, social media enables users to communicate asynchronously, something less possible with traditional landline or mobile phone communication. To put it simply, social media is faster, cheaper, more widespread, and more accessible than traditional forms of ICTs, and it is therefore significantly different than those traditional forms of communication (Tufekci & Wilson, 2012). Such significant differences between traditional ICTs and social media make it relevant to study the effect of social media specifically, particularly considering the fact that most previous studies did either not disaggregate social media from the Internet or are not based on a large-N analysis (Adhami, 2007; Bailard, 2015; Pierskalla & Hollenbach, 2013; Tähtinen, 2021; Weidmann, 2015)

Moreover, social media has also been demonstrated to facilitate collective action, leading some to coin the term "liberation technology," stressing its capacity to overcome collective action problems and "expand the horizon of freedom" (Diamond, 2010, p. 70). In line with this view, Steinert-Threlkeld et al. (2015) observed that coordination on social networks correlates with the onset of offline protest. Similarly, Tufekci & Wilson (2012) argued that social media facilitates protest through the possibility of reaching a great number of likeminded people. This line of thinking can also be applied to the onset of ethnic conflict, given the argument that social media access can also facilitate the collective action needed to engage in armed conflict.

The onset of intrastate conflict can be regarded as a collective problem, because fighting against the government can be very costly for the individual, while the benefits of potential power gains relative to the government can be enjoyed by a larger group. This pattern encourages free-riding (Olson, 1989; Pierskalla & Hollenbach, 2013). Social media allows group leaders to reward participation, while punishing free-riders. Due to the nature of social media, this can be done regardless of the geographical location of the group leaders and group members, as well as in a temporal independent context. Additionally, social media allows groups to coordinate more effectively due to the possibility to communicate from almost anywhere with each other, while also making it possible for groups to share information about their enemy in a secure environment (Matyasik, 2014). The mechanism by which social media can facilitate conflict onset are depicted in Figure 1.

Figure 1

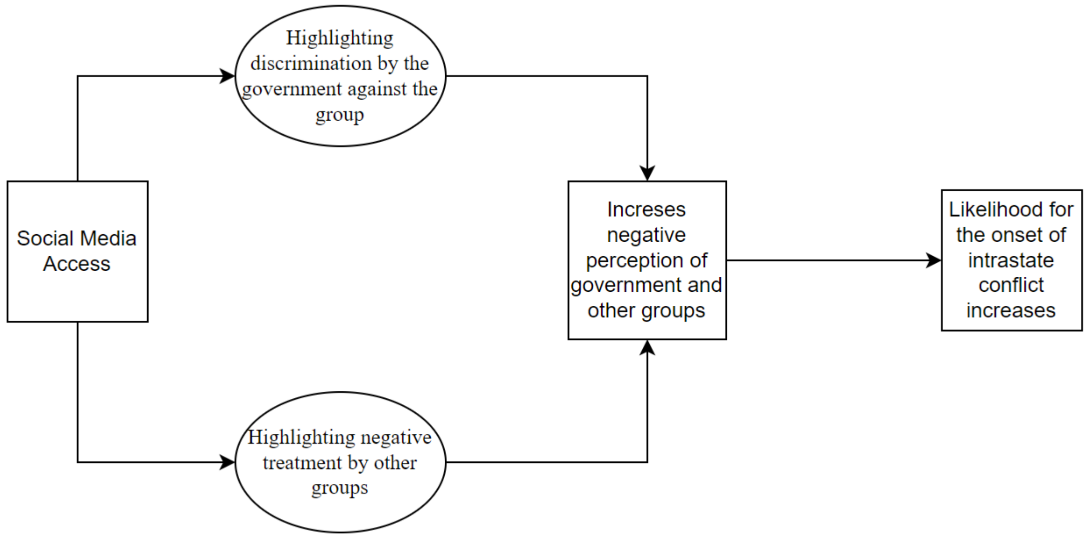
Mechanisms social media and collective action



Moreover, social media also highlights grievances of group members against other groups and the government. One of the master drivers for the onset of intrastate conflict are emotions (Tang, 2015). Negative emotions towards another group or the government can change the perception of that group, and the perception of the intentions of that group, negatively. If the intentions of one group are seen as detrimental for another group, intrastate conflict onset becomes more likely, due to a perceived need for self-defense (Tang, 2015). Additionally, if one's own group is perceived as significantly worse off than another group, or the group is perceived as discriminated against by the government, the potential benefits of insurgency increase, and intrastate conflict onset becomes more likely (Tang, 2015). Through social media these emotions towards another group or the government can be altered negatively through sharing information that depicts such discrimination (Sunstein, 2018). Therefore, social media can make the onset of intrastate more likely, by highlighting and showcasing discrimination and negative opinions towards other groups, increasing the perceived benefits of insurgency, and therefore, making intrastate conflict onset more likely. This mechanism is depicted in Figure 2.

Figure 2

Mechanisms social media and grievances



In light of the collective action and the grievance mechanisms, I derived the following hypothesis:

H1: An increase in access to social media will increase the likelihood for the onset of intrastate conflict.

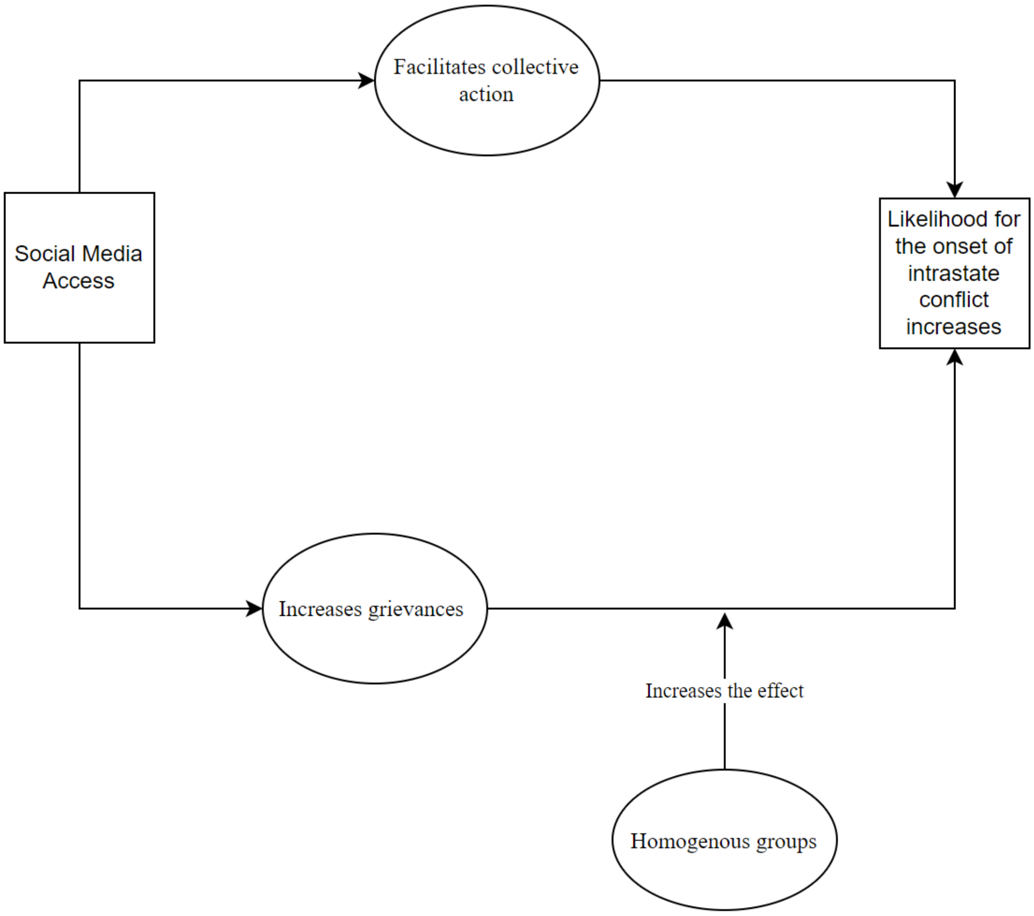
Furthermore, research has shown that conflict onset becomes more likely in environments in which ethnic fragmentation, and especially strong ethnic polarization, are present (Montalvo & Reynal-Querol, 2005; Sambanis, 2001). Combining these findings with the theorization by Tang (2015) leads to the following causal path. Social media access can increase polarization between groups, thereby negatively impacting the perception of other groups, the onset of conflict becomes more likely.

This follows Sunstein's (2018) observation that in homogenous groups, opinions become more extreme. Due to the design of social media platforms, echo chambers are created, increasing group fragmentation, and therefore group polarization. This, in turn, increases the likelihood for the onset for intrastate conflict. When groups only interact with their kin on social media, their opinions become more extreme. This increases the grievances that groups can have (Posen, 1993; Tang, 2015). If the intentions of outside groups are perceived to be hostile, the onset of ethnic conflict becomes more likely. Group polarization takes place only when social media platforms are divided into homogenous subgroups.

The mechanisms that link the homogeneity of a group to an increase in the likelihood for the onset of conflict are summarized in Figure 3.

Figure 3

Mechanisms hypothesis 2



Based on these mechanisms, the following second hypothesis is proposed:

H2: The effect of social media access on the onset of intrastate conflict is stronger in countries that are more homogenous.

4. Research Design

4.1 Case Selection and Models

The hypotheses discussed above were tested using a binomial logistic regression model with the onset of conflict as the dependent variable. To identify the effect the ethnic homogeneity of the country has on the effect size of social media access, some models included an interaction effect. The unit of analysis is a country-year dyad and covers the years from 2000 to 2020. The sample includes all countries listed in the GROW^{Up} dataset, where countries are defined as being administered by an intact sovereign state and having more than 250.000 inhabitants in the year of 1990 (Vogt et al., 2021). Importantly, this definition excludes overseas colonies, very small states, and failed states but newly independent states are included, starting from the year of their independence (Vogt et al., 2021). According to this definition, the dataset includes 3600 country-years observations for 173 unique countries.

Contrary to previously conducted research on the effects of the Internet on the onset of intrastate conflict, which focused mainly on cases studies or small-N comparisons, a large-N analysis will reduce sample bias to a minimum while allowing for more generalizable results (Adhami, 2007; Asimovic, Nagler, Bonneau, & Tucker, 2021; Barbera et al., 2015; Howard & Hussain, 2011; Tähtinen, 2021). Therefore, this research takes into account data from all countries in which data are available because, based on the theorization, a large difference between countries, regions, or even continents, warranting the exclusion of certain countries or country groups, should not exist. Additionally, the year 2000 has been chosen as the start year for the analysis due to the emergence and launches of major social media websites, such as MSN Messenger in 1999, LinkedIn and MySpace in 2003, and Facebook in 2004. The upper limit of 2020 was chosen due to the unavailability of more recent disaggregated data on conflict onset.

4.2 Variables

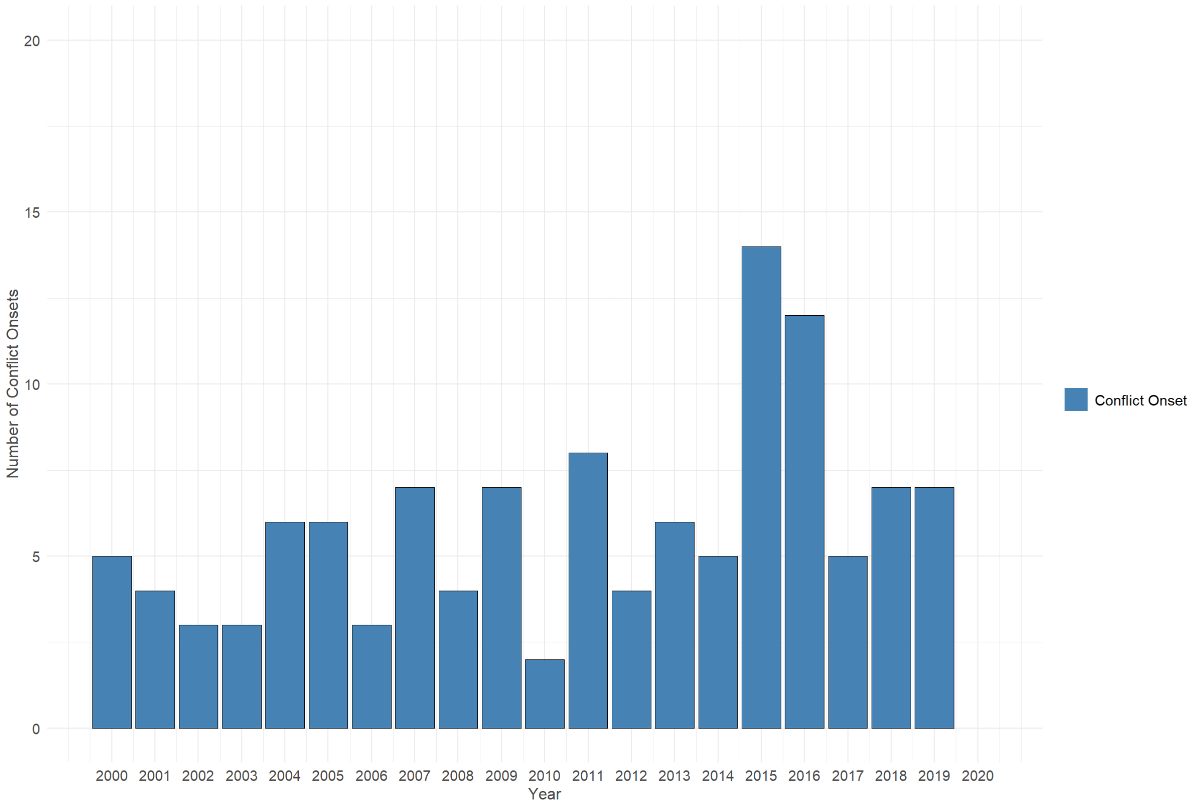
4.2.1 Dependent Variable

The dependent variable measures whether or not a country experienced the onset of intrastate conflict. Intrastate armed conflict is defined as a "conflict between a government and a non-governmental party" that resulted in at least 25 battle related deaths within one calendar year (Gleditsch et al., 2002, p. 618). If a country experienced the onset of war within a calendar year, and within the two preceding years, there was no active intrastate armed conflict, the variable was coded as 1; if there was no onset of conflict, the variable was coded as 0.

Over the 20-year period, from 2000 to 2020, 118 intrastate armed conflicts were reported for the 3633 country-years. Figure 4 shows the distribution of conflict onset over this 20-year time period. The peak of 14 intrastate conflict onsets was in 2015, whereas the year 2010 experienced the fewest conflict onsets within this 20-year period. This increase in 2015 was mainly driven by conflicts in Syria, Iraq, and Afghanistan, linked to Islamic extremism (Dupuy et al., 2016)

Figure 4:

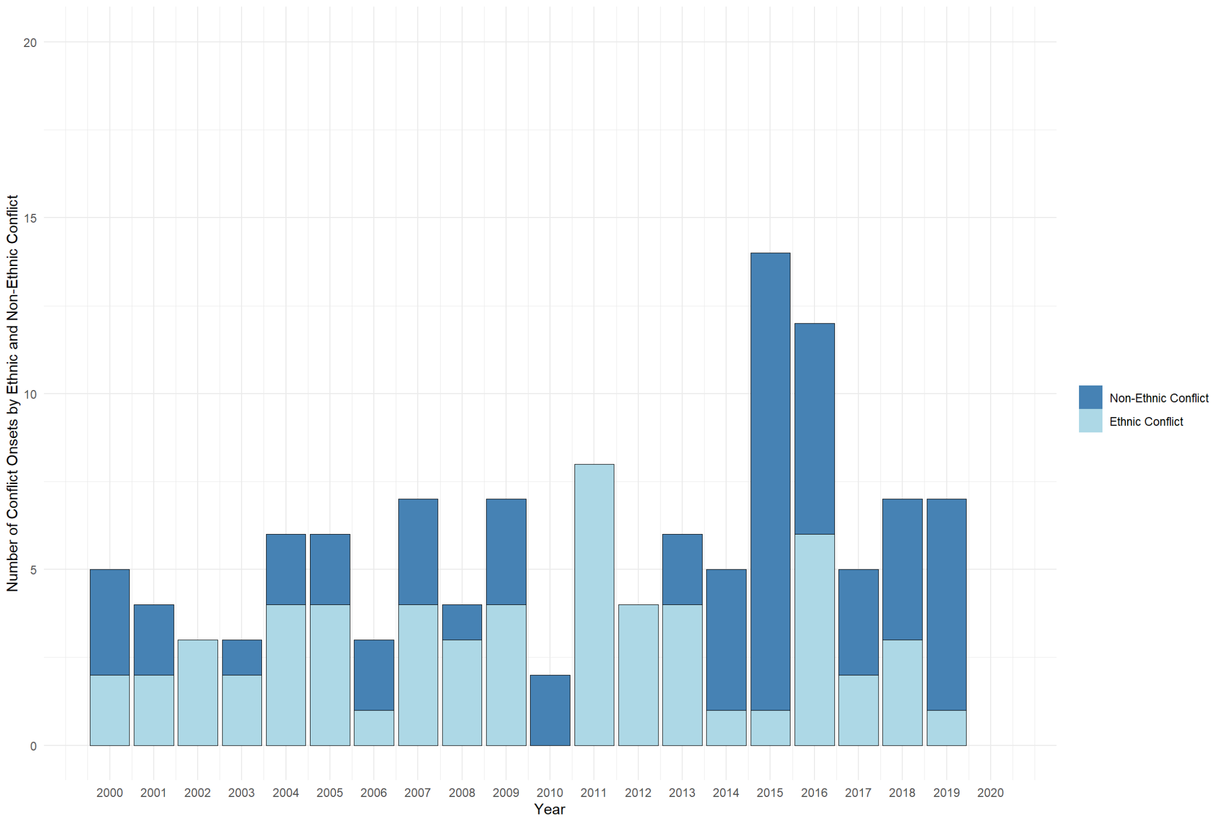
Distribution of Intrastate Conflict Onset



Additionally, the GROW[Up] dataset allows for the distinction between ethnic and non-ethnic intrastate conflict onset. A conflict is classified as an ethnic intrastate conflict if at least one ethnic group is linked to the conflict (Vogt et al., 2021). Of the total 118 intrastate armed conflicts, 58 were fought along ethnic lines, whereas 61 were non-ethnic conflict. The distribution of conflict onset separated by the two sub-categories is depicted in Figure 5. The year 2016 had the highest number of ethnic intrastate conflict onsets (6), and 2015 was the year with the highest number of non-ethnic conflict onsets (13). The high number of non-ethnic conflicts was again driven largely by Islamic extremism (Dupuy et al., 2016). Ethnic conflicts in the year 2016 were recorded in Tunisia, Philippines, Pakistan, Bangladesh, Syria, and Jordan. It is important to note there was no recorded intrastate conflict onset in the year 2020. Due to the corona virus, it is possible that the data for the year of 2020 is unreliable, or that the onset of the corona virus had grave impacts on the onset of intrastate conflict. Regardless, it is important to keep this outlier in mind while performing the logistic regression analysis, without providing a valid reason for the exclusion of the observations.

Figure 5:

Distribution of conflict onset separated by subgroup



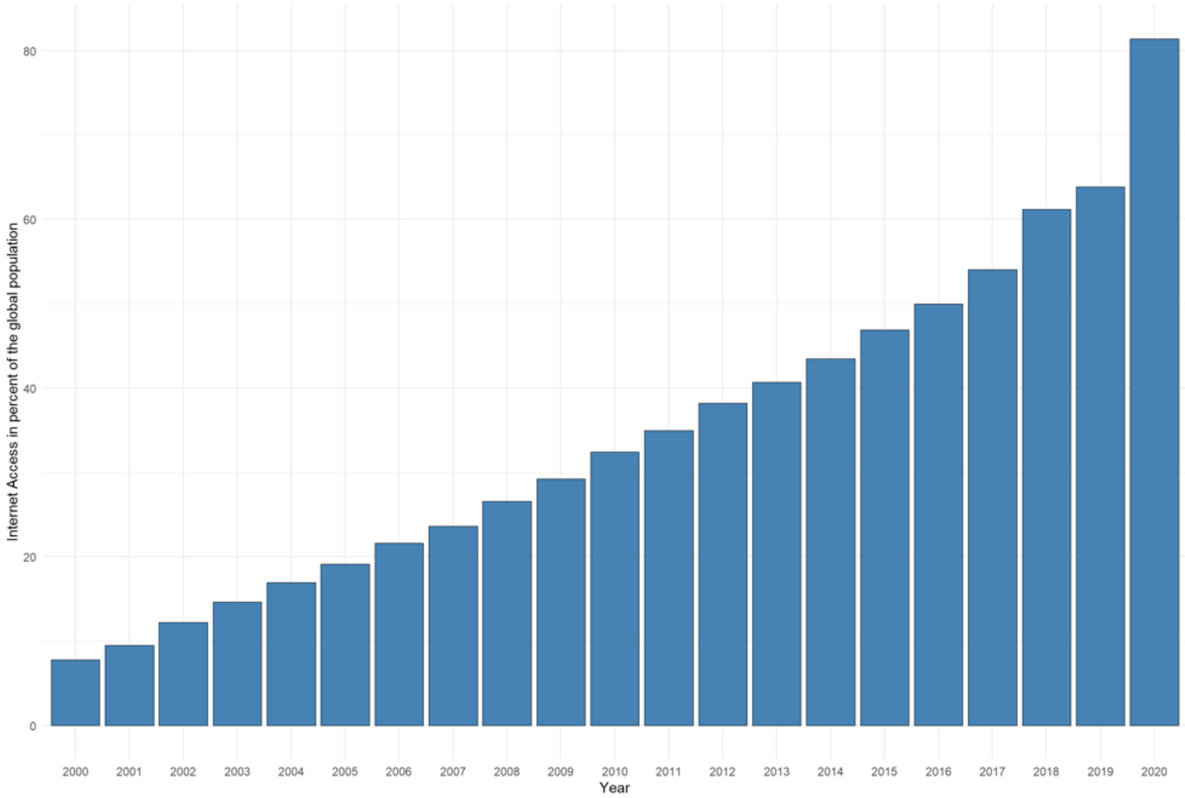
4.2.2 Independent Variables

Due to data limitations on the access to social media, this research used two different independent variables in order to estimate the true effect of social media access on the onset of conflict: Internet access and social media penetration. Internet access will be used as a proxy variable for the access to social media. Internet access is measured as share of the population within a country which has Internet access. People are counted as having Internet access if they have used the Internet in the last three months. This data is provided by the International Telecommunication Union (ITU) and published by the World Bank (World Bank, 2022a).

Figure 6 depicts the average of Internet access of all countries over the time period of 2000 to 2020. Unsurprisingly, 2000 is the year with the lowest amount of people with Internet access, in which only under 10% of the global population had access to the Internet, whereas 2020 is the year with the highest percentage of people having access to the Internet. The increase from year 2019 to 2020 is larger than any prior increase. This could be the case because only countries with higher Internet access rates reported their numbers, while countries with lower Internet access rates had not yet reported their numbers. This would increase the average Internet access rates but would not be an accurate representation of the true numbers. As the year 2020 is the year with the highest percentage of Internet access but with no recorded onset, this might bias the results of the statistical analysis towards underestimating the effects that Internet access has on the onset of conflict. To account for this potential bias, and the potential influence this bias could have on the analysis, models including and excluding 2020 were conducted.

Figure 6:

Average Rate of Internet Access



The second independent variable is social media penetration. Social media penetration is measured as the share of social media users as percentage of the whole population. The data used in this research was collected from reports by the marketing companies "We are Social" and "KEPIOS," which publish reports on digital media use for over 170 countries (Kemp, 2022). However, the bulk of these reports only start in the year 2017, with some occasional reports for the years 2011, 2015, and 2016. Therefore, using a Kalman[2] smoothing technique for the imputation of time-series data, the missing data between the years 2000-2017 was

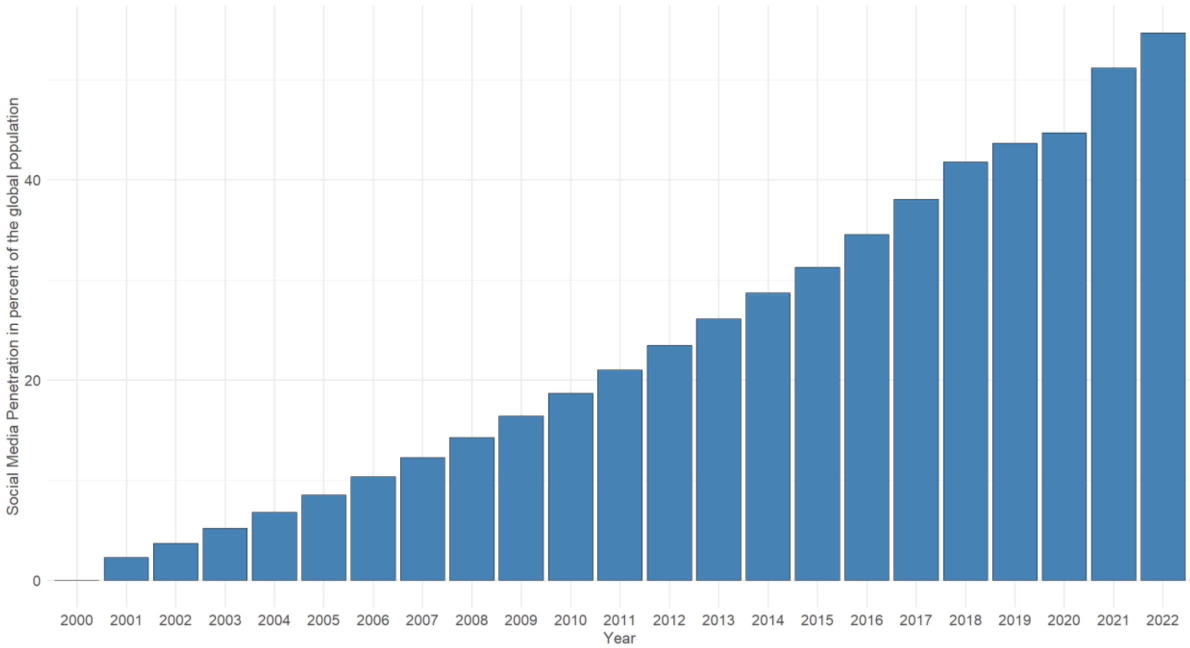
[2] "Kalman smoothing algorithm uses a series of measurements observed over time, containing noise and other inaccuracies, and produces estimates of unknown variables. This estimate tends to be more accurate than those based on a single measurement alone" (Maitra, 2021).

See appendix 1 for more information on how it was done specifically for this research.

imputed to match the time frame of the Internet access data. Figure 7 shows the average social media penetration per year, demonstrating that the imputed data is approximately linearly distributed. Unsurprisingly, this distribution is similar to the distribution of Internet access.

Figure 7:

Average Rate of Social Media Penetration



Both independent variables come with their own unique advantages and disadvantages. Access to the Internet is a very reliable measure that has been measured since the 1990s. However, it is not disaggregated enough to capture the potential effects of social media access on conflict onset, which are the focus of this research. Social media penetration is a very accurate measure for social media access; however, the data availability is limited to the years 2017-2022, and therefore, imputation is needed for the years 2000 to 2016 in order to make a meaningful analysis and comparison between the models possible but not showing the true values of social media penetration.

4.2.3 Control Variables

Additionally, a number of relevant control variables are included in the models. These control variables have been proven to be statistically significant in predicting the onset of intrastate armed conflict in previous studies or are commonly included as controls in analyses of conflict onset (Bailard, 2015; Paul Collier & Hoeffler, 2000, 2004; Denny & Walter, 2014; Hegre & Sambanis, 2006; Pierskalla & Hollenbach, 2013; Sambanis & Schulhofer-Wohl, 2019, 2019; Weidmann, 2015; Weidmann & Rød, 2019). These include GDP per capita or annual GDP growth, percentage of young male population, percentage of rural population, military expenditure, population, and population growth. The data on these control variables was collected by the World Bank and is available in the World Bank databank (World Bank, 2022b). To control for the effects of regime type on the onset of intrastate war, the categorization of the V-Dem project of countries into closed autocracies, electoral autocracies, electoral democracies, and liberal democracies, is included in the analysis (Coppedge et al., 2019). The ruggedness/percentage of mountainous terrain of a country also influences the onset of intrastate war. Therefore, a ruggedness index is included in the analysis (Riley, DeGloria, & Elliot, 1999). The descriptive statistics of all dependent, independent, and control variables are depicted in Table 1.

Table 1: Descriptive Statistics for all variables of Model 1, 2, and 3 (dependent and independent)

	N	Minimum	Maximum	Mean	Std. Deviation
Onset of intrastate Conflict	3600	0.00	1.00	0.03	0.178
Onset of ethnic intrastate conflict	3600	0.00	1.00	0.02	0.127
Onset of non-ethnic intrastate conflict	3600	0.00	1.00	0.02	0.131
Internet Access	3298	0.00	100.00	32.46	30.57
Social Media Penetration (Imputed)	3556	0.00	116.00	23.38	24.28
Regime Type	3559	0.00	3.00	1.61	0.99
GDP per Capita PPP	3495	435.08	141634.71	17040.64	19654.73
Military expenditures	3190	0.00	778232000000.00	9467551847.90	52077888017.37
GDP Growth Annual	3536	-62.08	123.14	3.54	5.56
Population	3633	247310.00	1410929362.00	39819152.71	141870174.53
Rural Population %	3603	0.00	91.75	43.34	22.60
Rugged Terrain	3570	0.01	6.74	1.31	1.20
Population Growth Rate	3624	-4.53	17.51	1.50	1.52
Natural Resources Rent	3540	0.00	87.58	8.07	11.85
Young Male Population %	3594	10.41	25.00	17.76	3.04

5. Analysis

5.1 Results

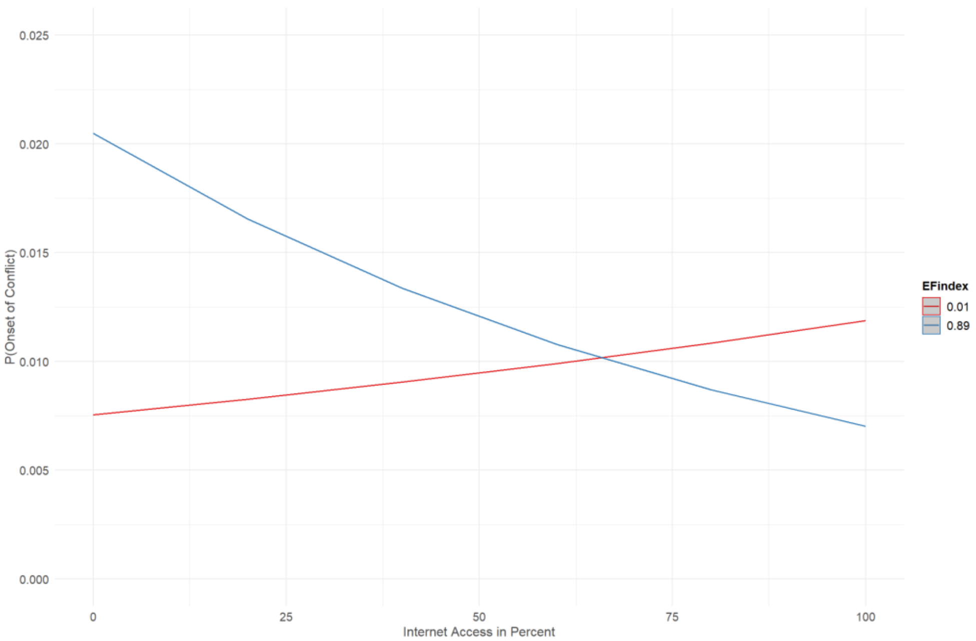
Tables 2 and 3 present the results of the logistic regression analysis. Model 1 is a general model for intrastate conflict onset in general, whereas model 2 focuses solely on ethnic intrastate conflict, and model 3 on non-ethnic intrastate conflict. In order to test hypothesis 2, models 4 and 8 include an interaction effect between internet access and ethnic fractionalization in order to establish whether or not the effect changes depending on the ethnic fractionalization of the country. Model 5, 6, 7, and 8 are the same models with social media penetration as independent variable instead of Internet access.

In model 1, Internet access has an odds ratio of 1.014, meaning that with a one percent increase in the access to the Internet, intrastate conflict onset becomes slightly more likely. If 100 percent of the population has access to the Internet, intrastate conflict onset becomes roughly 10 percent more likely, ceteris paribus. However, this effect is not statistically significant at any conventional significance level. Likewise, as shown in model 2, Internet access is also slightly positively correlated with the onset of ethnic intrastate conflict. With an odds ratio of 1.006, the effect of Internet access on the onset of intrastate ethnic conflict is very small. Additionally, this effect is also not statistically significant. Interestingly, model 3 predicts that a one percent increase in Internet access increases the odds for the onset of non-ethnic conflict by 1.024. This effect is statistically significant at the 0.05 significance level.

Moreover, model 4 shows that when including an interaction effect between Internet access and ethnic fractionalization, Internet access is still positively correlated with the onset of intrastate conflict. This is not statistically significant at any conventional significance level. Additionally, the interaction effect between Internet access and ethnic fractionalization demonstrates that in countries with higher ethnic fractionalization, higher Internet access decreases the odds for the onset of intrastate conflict. This effect is also not statistically significant at any conventional significance level. This interaction effect is visualized in Figure 8. For easier understanding, the two most extreme values that are present in the dataset of the ethnic fractionalization index have been chosen, 0.01 in North Korea in the year 1990, and 0.89 in Liberia in the year 1990. As shown in the plot of the interaction effect, the probability for the onset of intrastate conflict decreases with an increase in Internet access for an ethnic fractionalization value of 0.89, while there is an increase for an ethnic fractionalization value of 0.01.

Figure 8:

Interaction Effect between Internet Access and Ethnic Fractionalization

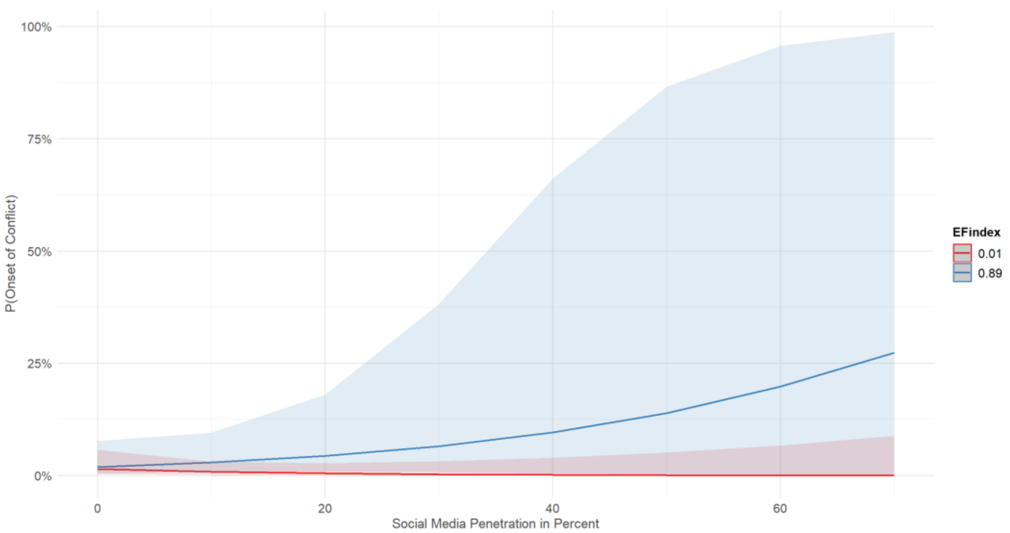


Lastly, Nagelkerke's R^2 and Cox and Snell's R^2 are relatively small for all four models, meaning that the explanatory power of the models is limited but still acceptable (Peng, Lee, & Ingersoll, 2002). It is important to note that unlike in OLS-regression, Nagelkerke's R^2 and Cox and Snell's R^2 do not measure the variance that is explained by the model. Nonetheless, the increase in the -2LL ratio indicates that especially model two and model three are better fitted for the data than model one. This comparison is not meaningful with model four due to the addition of extra independent variables because the addition of extra variables will automatically increase the -2LL ratio.

The results of the analysis with social media penetration as independent variable are summarized in Table 3. Generally, the results between social media penetration and Internet access are similar, but different in two key aspects. [3] First, while the effects of social media penetration on conflict onset are similar to the effects of Internet access, the effect of social media is not statistically significant for any type of conflict. Secondly, the interaction effect between social media penetration and the ethnic fractionalization index is positive, indicating that in countries with higher ethnic fractionalization, social media has a bigger effect on the onset of conflict, compared to countries with lower ethnic fractionalization as depicted in Figure 9. However, this difference is not statistically significant.

Figure 9

Interaction Effect between Social Media Penetration and Ethnic Fractionalization (95% CI)



[3] It is important to note that running the analysis on the original data without imputation leads to different results. The results of the original data are found in appendix 3.

Table 2: Logistic Regression Results: How does Internet Access influences the Onset of Intrastate Conflict (Odds Ratio with 95%CI Interval)

	Intrastate conflict onset (Model 1)	Ethnic intrastate onset (Model 2)	Non-ethnic intrastate onset (Model 3)	Interaction effect (Model 4)
Internet Access %	1.014	1.006	1.025*	1.005
	[0.998, 1.030]	[0.983, 1.029]	[1.004, 1.046]	[0.957, 1.055]
Regime Type: [Ref. = Closed Autocracy]				
Electoral Autocracy	5.252**	3.049	8.050*	1.724
	[1.867, 14.772]	[0.949, 9.794]	[1.480, 43.771]	[0.572, 5.195]
Electoral Democracy	2.402	1.315	3.148	0.793
	[0.849, 6.799]	[0.382, 4.526]	[0.595, 16.649]	[0.210, 3.004]
Liberal Democracy	1.489	1.044	1.344	0.602
	[0.348, 6.366]	[0.155, 7.005]	[0.136, 13.323]	[0.089, 4.076]
GDP per capita PPP	1.000**	1.000*	1.000	1.000*
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Military expenditure	1.000	1.000	1.000	1.000*
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
GDP Growth Annual	1.039	1.042	0.987	1.069*
	[0.991, 1.090]	[0.989, 1.098]	[0.913, 1.067]	[1.015, 1.127]
Population	1.000***	1.000***	1.000**	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rural Population	0.990	0.984	0.999	0.985
	[0.975, 1.005]	[0.964, 1.004]	[0.978, 1.021]	[0.963, 1.007]
Rugged Terrain	1.079	1.235	0.854	1.263
	[0.880, 1.322]	[0.936, 1.631]	[0.636, 1.146]	[0.938, 1.702]
Pop-Growth Rate	1.503***	1.302	1.717***	1.294
	[1.210, 1.866]	[0.943, 1.797]	[1.288, 2.290]	[0.930, 1.801]
Natural Resources Rent	1.013	1.040**	0.957*	1.023
	[0.992, 1.034]	[1.015, 1.065]	[0.917, 0.999]	[0.995, 1.052]
Young Male Population	0.976	0.889	1.074	0.827
	[0.836, 1.140]	[0.727, 1.087]	[0.855, 1.350]	[0.662, 1.032]
Ethnic Fractionalization [EF]				3.157
				[0.485, 20.536]
Interaction - Internet Access:EF				0.983
				[0.890, 1.084]
(Constant)	0.011**	0.076	0.001**	0.384
	[0.001, 0.230]	[0.002, 3.191]	[0.00001, 0.066]	[0.005, 28.927]
-2LL	-341.68	-202.66	-202.86	-171.93
Cox and Snell's R^2	0.04	0.03	0.02	0.03
Nagelkerke's R^2	0.15	0.16	0.14	0.14
N	2,819	2,819	2,819	1,851

Note: Odds ratios with 95% confidence intervals in brackets.
****p < 0.001, **p < 0.01, * p < 0.05*

Table 3: Logistic Regression Results: How does Social Media Penetration (Imputed Data) influences the Onset of Intrastate Conflict (Odds Ratio with 95%CI Interval)

	Intrastate conflict onset (Model 5)	Ethnic intrastate onset (Model 6)	Non-ethnic intrastate onset (Model 7)	Interaction Model (Model 8)
Social Media Penetration	1.009	1.003	1.017	0.945
	[0.992, 1.027]	[0.977, 1.030]	[0.994, 1.041]	[0.865, 1.033]
Regime Type: [Ref. = Closed Autocracy]				
Electoral Autocracy	2.618*	2.725	2.068	1.167
	[1.109, 6.184]	[0.865, 8.585]	[0.655, 6.528]	[0.436, 3.128]
Electoral Democracy	1.337	1.200	1.010	0.529
	[0.551, 3.244]	[0.356, 4.038]	[0.314, 3.249]	[0.155, 1.807]
Liberal Democracy	0.912	1.000	0.505	0.476
	[0.231, 3.594]	[0.147, 6.793]	[0.069, 3.667]	[0.080, 2.833]
GDP per Capita PPP	1.000**	1.000*	1.000	1.000**
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Military expenditure	1.000	1.000	1.000	1.000*
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
GDP Growth Annual	1.031*	1.025	1.025	1.040*
	[1.003, 1.060]	[0.994, 1.057]	[0.964, 1.089]	[1.007, 1.075]
Population	1.000***	1.000***	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rural Population	0.991	0.985	0.999	0.984
	[0.976, 1.006]	[0.965, 1.006]	[0.979, 1.019]	[0.961, 1.008]
Rugged Terraine	1.068	1.251	0.843	1.265
	[0.877, 1.301]	[0.956, 1.639]	[0.630, 1.128]	[0.941, 1.702]
Pop Growth Rate	1.429***	1.286	1.503**	1.363
	[1.159, 1.763]	[0.931, 1.775]	[1.141, 1.980]	[0.984, 1.890]
Natural Resources Rent	1.008	1.038**	0.958*	1.021
	[0.988, 1.028]	[1.015, 1.062]	[0.921, 0.996]	[0.993, 1.050]
YoungMalePopulation15to24	0.973	0.876	1.072	0.808
	[0.837, 1.131]	[0.717, 1.070]	[0.864, 1.330]	[0.646, 1.010]
EFindex				1.320
				[0.215, 8.088]
Social-Media-Penetration:EFindex				1.117
				[0.950, 1.313]
(Constant)	0.028*	0.112	0.003**	1.739
	[0.002, 0.478]	[0.003, 3.991]	[0.00005, 0.217]	[0.027, 112.407]
-2LL	-372.67	-211.49	-230.58	-176.96
Cox and Snell's R^2	0.032	0.023	0.019	0.026
Nagelkerke's R^2	0.13	0.15	0.12	0.13
N	3,035	3,035	3,035	1,891

Note: Odds ratios with 95% confidence intervals in brackets.
****p < 0.001, **p < 0.01, * p < 0.05*

5. 2 Interpretation of the results

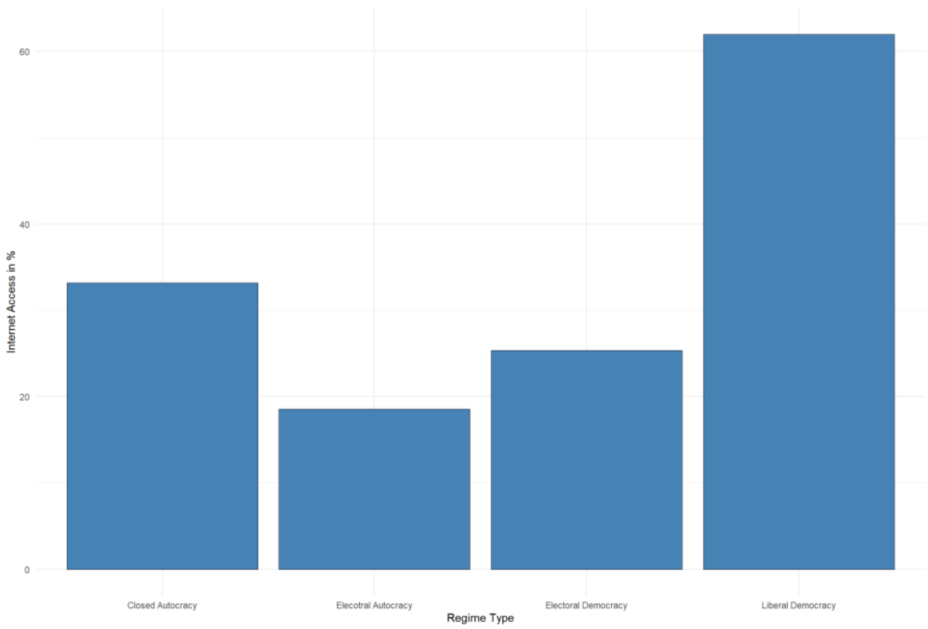
5.2.1 Hypothesis 1

Based on the results of model 1 and model 5, hypothesis 1 stating that an increase in social media access leads to a higher likelihood of intrastate conflict onset is not supported. This could have several reasons. First, and most obviously, it is possible that social media access alone is not sufficient in reducing the problems that are associated with collective action and could even aggravate them. Recent studies have shown the Internet and social media in particular can also be used by governments in order to make collective action less likely and grievances less visible. For example, King et al. (2014) showed that the Chinese government uses posts on social media platforms actively to distract from topics that have the potential to lead to collective action. Since there is a diffusion of repression technologies between autocratic states, it is likely that other autocratic and hybrid regimes employ similar measures (Olar, 2019). Additionally, Gunitsky et al. (2015) demonstrated that autocratic rulers can use social media to make collective action less likely and grievances less visible, by using social media as a tool for counter-mobilization, framing discourse, preference divulgence, and elite coordination. However, this is contrary to the findings in both model 1 and 5, which indicate that electoral autocracies are more likely to experience the onset of conflict. Therefore, another explanation is more likely.

Second, it is possible there is a threshold of social media users that needs to be reached before collective action can effectively be organized and enable grievances to spread to a large group of people. Figure 10 shows the average Internet access in percentage by regime type according to the categorization of the V-Dem index (Coppedge et al., 2019). Especially in electoral democracies and electoral autocracies, so called hybrid regimes, people have less access to the Internet than in liberal democracies and closed autocracies (Figure 10). However, hybrid regimes are also the regimes that are most prone for the onset of intrastate conflict (Stockemer, 2010). This is also supported by the statistically significant odds ratio of 5.252 for electoral autocracies (model 1), meaning that intrastate conflict onset in electoral autocracies is more than 5 times as likely as in closed autocracies. Therefore, it could be that social media access does indeed lead to the onset of intrastate conflict, but the effect is not yet visible due to the low number of people that have access to social media in hybrid regimes. Since social media access is also increasing in hybrid regimes, it could be beneficial to revisit this research in a couple of years, when a larger number of people have access to social media in hybrid regimes.

Figure 10

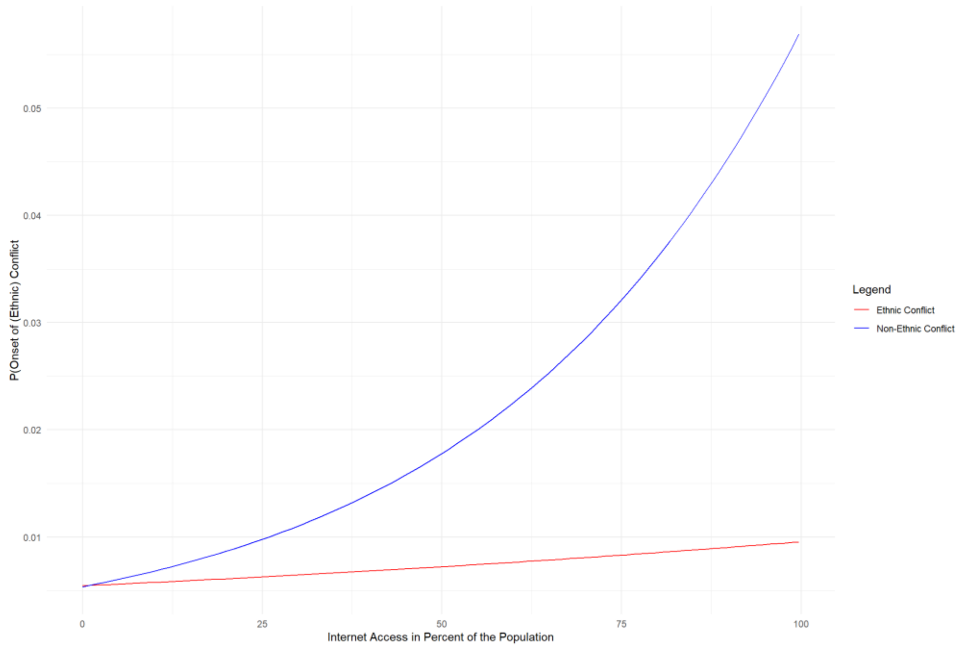
Internet Access by Regime Type



Third, it could be the case that intrastate conflict onset is too broad of a category. It is possible that social media access only leads to the onset of some types of intrastate conflict. This is supported by models 2 and 3 that disaggregate conflict into ethnic and non-ethnic conflict. Running the same models for ethnic intrastate conflict and non-ethnic intrastate conflict separately reveals that internet access is statistically significant at the 0.05 significance level for the onset of non-ethnic intrastate conflict, while it is not statistically significant for the onset of ethnic intrastate conflict. This difference between ethnic intrastate conflict and non-ethnic intrastate conflict is depicted in Figure 11, where it is clearly visible there is a significant increase in the likelihood of the onset of intrastate war for non-ethnic conflict, whereas there is no such increase for ethnic conflict. It is important to note that while showing a similar effect, the effect is not statistically significant for the models based on social media penetration as the independent variable. The effect that Internet access has on the onset of non-ethnic conflict is more than five times the effect that Internet access has on ethnic conflict, providing some evidence that suggests the two subcategories of intrastate conflict should be analyzed separately. There is a number of possible explanations as to what could lead to this difference between ethnic and non-ethnic conflict.

Figure 11

Comparison Likelihood Onset of Conflict Ethnic and Non-Ethnic Intrastate Conflict



Based on the theorization within this thesis, social media access leads to an increase in the likelihood of the onset of intrastate conflict because it facilitates collective action and increases the visibility of grievances. The difference that access to social media has on the onset of non-ethnic intrastate conflict compared to ethnic intrastate conflict could be explained by the possibility that in ethnic contexts, social media is not (or is less) needed to make grievance's more visible and to facilitate collective action within an ethnic group.

As Denny and Walter (2014) and Weidmann (2009) argued groups that are organized around ethnicity are better able to overcome collective action problems due to geographical proximity and increased ease of communication. Additionally, group members often have long-lasting grievances against other groups due to historical power imbalances or active discrimination of the state against the ethnic group (Cederman, Weidmann, & Gleditsch, 2011; Horowitz, 1985). These two factors are assumed to make the onset of ethnic intrastate conflict more likely.

According to the theorization within this thesis, social media access fulfills a similar role in facilitating collective action and highlighting grievances and making them more visible, similar to ethnicity. Therefore, it is possible that internet access does not have a statistically significant effect on the onset of ethnic intrastate conflict because the role of social media access to increase the likelihood of the onset of ethnic intrastate conflict is already taken on by ethnicity itself, making the effect that social media access has on the onset of ethnic intrastate conflict less strong and therefore statistically insignificant.

However, these findings also contradict the logic proposed by Weidmann (2015). He argued that ethnic conflict can spread through telecommunication networks by shedding light on grievances in other countries and demonstrates the successfulness of ethnic rebel movements as well as share information and strategy. This sharing of information can lead to a more successful insurgence, reducing potential opportunity costs, and therefore making ethnic intrastate conflict onset more likely. As social media access is working in a similar manner as telecommunication networks, following Weidmann's (2015) argument, it would have been logical that social media access would lead to an increase in the likelihood of the onset of intrastate ethnic conflict. Therefore, it is possible that the mechanisms discussed in the theory section only apply to non-ethnic conflict.

5. 2. 2 Hypothesis 2

The lack of statistically significant results in the interaction model (model 4 and model 8), leads to the rejection of hypothesis 2. Moreover, model 4 and model 8, while both not statistically significant, show different effect sizes. While model 4 predicts that ethnic homogeneity leads to an increased effect of Internet access on the onset of conflict, model 8 predicts the opposite. This demonstrates again the need to disaggregate social media from the Internet in the analysis of their effect on conflict onset.

The implications of these finding are clear. As theorized, high network homogeneity makes the effect of social media access stronger as group opinions become more and more polarized. Therefore, the insignificant results from model 4 and 8 indicate that group polarization does not impact the effect that social media access has on the onset of intrastate conflict. This could have several explanations. First, and most obviously, country homogeneity simply does not matter in the real world when it comes to the effects of social media on the onset of intrastate conflict, either because group polarization does not matter for the onset of intrastate conflict, or because group polarization does not take place on social media platform. If group polarization does not happen on social media platforms, this would contradict a large number of previous finding (Asimovic et al., 2021; Iandoli, Primario, & Zollo, 2021; Yardi & Boyd, 2010). For example, Guerra et al. (2013) found that group polarization happens in several scenarios, including political debates such as same-sex marriage, abortion, and gun control. Therefore, it would be inconsistent if group polarization would not also happen on topics such as secession, ethnic group discrimination, and insurgency against the government. Though, it is possible that such discussions do not take place in public forums, but in already separated and private groups within social media.

Second, and more likely, it is possible that ethnic fractionalization is not a good proxy measure for the homogeneity of a social media platform, especially considering that social platforms are often fragmented into a lot of different subcategories. These subcategories could be a highly diverse and unrepresentative sample of the country's population. Simply put, a country being homogenous does not automatically make the people that use social media platforms also homogenous or vice versa. This effect could drastically alter the results of the interaction analysis. Consequently, a measure that could directly gauge the homogeneity of a social media platform could lead to different results.

5. 3 Robustness

To test the robustness of the above discussed results, several additional models were run[4]. Some models include a logarithmic transformation of GDP per capita PPP and population. A logarithmic transformation is usually done for variables that do not satisfy the assumption of linearity. However, often, performing a logarithmic transformation of the variables can make them even more skewed. Interestingly, the models with a logarithmic transformation show roughly the same effect of Internet access on the onset of intrastate conflict, that was observed in the original models. However, transforming population and GDP per capita PPP logarithmically does change the significance of Internet access on the onset of non-ethnic conflict in the Internet access models.

Additional models only take data into account from 2005 to 2020. The year 2005 could be another possible cutoff point, because this is the year when one of the biggest social media platforms, Facebook, was launched. To make sure that the observed effect in the original models is not simply a one-off effect; it is important to compare it to other time periods. While it is still observable that the effect of Internet access is largest on the onset of non-ethnic

[4] The results of the additional models are depicted in tables A and B, found in appendix 2.

conflict, and lowest for the onset of ethnic intrastate conflict, all three odds ratios are not statistically significant. This could mean the observed statistically significant effects of the original models are simply based on a luckily chosen time period. Though, it could also stem from the fact that reducing the observed time period by five more years, makes the onset of intrastate conflict even less likely, and therefore, predictors less significant.

Lastly, due to the anomalies overserved in the data from the year 2020, some models exclude 2020 from the data. As expected, removing the data from 2020 from the model increases the effect that Internet access and social media has on the onset of intrastate conflict, ethnic intrastate conflict, and non-ethnic intrastate conflict. This increase could occur because there are no conflict onsets recorded for the year 2020, but it is simultaneously the year with the highest percentage of Internet access. While removing the year 2020 from the data completely must be strongly justified, models excluding the year 2020 give insight the data from 2020 does bias the results.

Lastly, other models were run including a measure for the availability of alternative information. Including a variable that measures this is important in order to distinguish whether the observed effects in the original models are truly due to an increase in social media access or if it is just the access to non-censored or less biased information. While the odds ratios of the models including the alternative information index are all not statistically significant, a similar pattern compared to Internet access and social media penetration is observable. This is evidence that the increase in conflict onset is not simply due to an increase in Internet access but also that the availability of uncensored information could play a role. However, it is important to note that a direct comparison between the models is not directly meaningful due to the different sample size caused by the lack of reliable data for the alternative forms of information index.

5. 4 Limitations

As with all research, this research comes with a number of limitations. I identify four main limitations, mainly concerned with the quality of data: the fact that conflict onset is a rather rare event, problems with data limitations on an accurate measurement for social media penetration, general problems with conflict data, and the existence of a large number of missing values.

Intrastate conflict is a rather rare event. Over the course of the 20-year period, of the total of 3600 observations, only 118 were recorded as intrastate conflict onset, which is roughly only 3.2% of the observations where intrastate conflict did occur. As King and Zeng (2001, p. 138) noted, logistic regression analysis applied to rare event data can "sharply underestimate the probability of rare events." This could introduce a strong bias into the models. Nevertheless, because adjusting for this problem would most likely increase the effect size and the statistically significant of the predictor variables, it is not such a problem for this statistical analysis.

A much more important limitation is that using a proxy for the measurement of another variable can never truly accurately measure the value of the variable that the research aims to measure. This is most likely the case when using Internet access in order to measure the access to social media. While there is certainly a strong correlation between the two, there could be country-level differences in the data. For example, in Myanmar Facebook is almost synonymous to the Internet, meaning that almost everyone with Internet access is most likely also using a social media website. There are other countries, such as China, where social media access is heavily regulated, and Internet access probably differs significantly from social media access. Simply, just because one has access to the Internet does not mean they also have an Instagram, Facebook, or Twitter account. Thus, using social media data is important, however, that data is only very limitedly available and imputation is needed, potentially biasing the data and not depicting the real values in social media access. For example, the imputed data for social media penetration in this research overestimates the rate of social media penetration. [5]

Additionally, while there are 3600 observations for intrastate conflict onset, models 1, 2, and 3 only take 2819 observations into account due to missing data, meaning that almost 800 observations had to be excluded from the models. Model 4 takes even fewer observations into account, as the ethnic fractionalization index is only available through the year of 2013. This exclusion due to missing values can introduce a serious bias into the data, because it is to be assumed that a country that experiences the onset of intrastate conflict also experiences difficulties gathering statistics on its population, such as GDP per capita or the amount of people living in rural areas. Luckily, for the utilized dataset, the proportion between conflict onset and no conflict onset remains roughly similar at around 3.2%. Additionally, another possible introduction for bias in the data is how the data is gathered. Since most data on conflict is gathered based on media reports, which could be faulty and also impact conflict onset itself, analysis where this is not controlled should be taken with a grain of salt (Weidmann, 2016).

6. Conclusion and Future Research

This research investigated the relationship of social media access and the onset of intrastate conflict. It was expected that an increase in social media access increases the likelihood for the onset of intrastate conflict and that this effect is stronger in countries that are more homogenous. The results indicate there is no general effect between social media access and the onset of intrastate conflict. However, the research also highlighted that Internet access does indeed increase the likelihood of the onset for non-ethnic intrastate conflict. Additionally, this research also demonstrated the need for the disaggregation between social media and the Internet in

[5] See appendix 1 for more information.

general due to different results between the models based on Internet access and social media penetration.

These findings advance the study of the onset of conflict in three ways. First, the findings based on Internet access as an independent variable show that an increase in social media access does increase the likelihood for the onset of non-ethnic intrastate conflict, ceteris paribus. Second, demonstrating that ethnic and non-ethnic conflict should be analyzed separately from each other indicates the drivers of ethnic intrastate conflict and non-ethnic intrastate conflict are different from each other. Third, it is also demonstrated that social media and the Internet should not be used synonymously, and disaggregation between the two is needed. Therefore, more extensive data on social media access is needed to draw more meaningful inference on its effects.

Additionally, while this research theorized a causal path through which access to social media can affect the onset of intrastate conflict, large statistical models are ill-suited to properly identify causal mechanisms. Therefore, further study would be beneficial to identify causal mechanisms and then test those systematically. Moreover, the difference in effect between the onset of non-ethnic and ethnic intrastate conflict could also stem from a difference in the use of social media. Consequently, it would be insightful to identify differences in the usage of social media between different groups. While there is the need for some additional research, this thesis provides evidence for the need in differentiation between non-ethnic and ethnic conflict and the differentiation between social media and the Internet in general, advancing conflict research and research on the effects of social media.

7. References

Adhami, W. (2007). The strategic importance of the Internet for armed insurgent groups in modern warfare. *International Review of the Red Cross*, *89*(868), 857–878.

Asimovic, N., Nagler, J., Bonneau, R., & Tucker, J. A. (2021). Testing the effects of Facebook usage in an ethnically polarized setting. *Proceedings of the National Academy of Sciences—PNAS*, *118*(25), 1–9. National Academy of Sciences.

Bailard, C. S. (2015). Ethnic conflict goes mobile: Mobile technology's effect on the opportunities and motivations for violent collective action. *Journal of peace research*, *52*(3), 323–337.

Barbera, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & Gonzalez-Bailon, S. (2015). The Critical Periphery in the Growth of Social Protests. *PloS one*, *10*(11), 1-15.

Bhavnani, R., & Miodownik, D. (2009). Ethnic Polarization, Ethnic Salience, and Civil War. *Journal of Conflict Resolution*, *53*(1), 30–49.

Blimes, R. J. (2006). The Indirect Effect of Ethnic Heterogeneity on the Likelihood of Civil War Onset. *Journal of Conflict Resolution*, *50*(4), 536–547.

Cederman, L.-E., Weidmann, N. B., & Gleditsch, K. S. (2011). Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison. *American Political Science Review*, *105*(3), 478–495.

Collier, Paul, & Hoeffler, A. (2000). *Greed and Grievance in Civil War*. Washington, DC: World Bank.

Collier, Paul, & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563–595.

Collier, Paul, Hoeffler, A., & Rohner, D. (2008). Beyond greed and grievance: Feasibility and civil war. *Oxford Economic Papers*, *61*(1), 1–27.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., et al. (2019). *V-Dem Codebook V9* (SSRN Scholarly Paper No. ID 3441060). Rochester, NY: Social Science Research Network. Retrieved April 1, 2022, from https://papers.ssrn.com/abstract=3441060

Denny, E. K., & Walter, B. F. (2014). Ethnicity and civil war. *Journal of Peace Research*, *51*(2), 199–212.

Diamond, L. J. (2010). Liberation Technology. *Journal of democracy*, *21*(3), 69–83.

Dupuy, K., Gates, S., Nygård, H. M., Rudolfsen, I., Strand, H., & Urdal, H. (2016). Trends in Armed Conflict 1946-2015. *Conflict Trends*, *8*(1), 1–4.

Fearon, J. D., & Laitin, D. D. (2000). Violence and the Social Construction of Ethnic Identity. *International organization*, *54*(4), 845–877.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, *97*(01), 75–90.

Fink, C. (2018). Dangerous Speech, Anti-Muslim Violence, and Facebook in Myanmar. *Journal of international affairs (New York)*, *71*(1.5), 43–52.

Gagnon Jr, V. P. (1995). Ethnic conflict as an intra-group phenomenon: A preliminary framework. *Revija za sociologiju*, *26*(1–2), 81–90.

Gleditsch, N. P., Wallensteen, P., Erikson, M., Sollenberg, M., & Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research*, *39*(5), 615–637.

Guerra, P., Jr, W. M., Cardie, C., & Kleinberg, R. (2013). A Measure of Polarization on Social Media Networks Based on Community Boundaries. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), 215–224.

Gunitsky, S. (2015). Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability. *Perspectives on politics*, *13*(1), 42–54.

Gurr, T. R. (1993). Why minorities rebel: A global analysis of communal mobilization and conflict since 1945. *International Political Science Review*, *14*(2), 161–201.

Hegre, H., & Sambanis, N. (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *The Journal of Conflict Resolution*, *50*(4), 508–535.

Horowitz, D. L. (1985). *Ethnic groups in conflict*. Berkeley [etc.]: University of California Press.

Howard, P. N., & Hussain, M. M. (2011). The Upheavals in Egypt and Tunisia: The Role of Digital Media. *Journal of Democracy*, *22*(3), 35–48.

Humphreys, M., & Weinstein, J. M. (2008). Who Fights? The Determinants of Participation in Civil War. *American Journal of Political Science*, *52*(2), 436–455.

Iandoli, L., Primario, S., & Zollo, G. (2021). The impact of group polarization on the quality of online debate in social media: A systematic literature review. *Technological Forecasting and Social Change*, *13*(4), 12-35.

Jennings, D. E. (1986). Outliers and Residual Distributions in Logistic Regression. *Journal of the American Statistical Association*, *81*(396), 987–990.

Kemp, S. (2022). Report Library. *DataReportal – Global Digital Insights*. Retrieved June 3, 2022, from https://datareportal.com/library

King, G., Pan, J., & Roberts, M. E. (2014). Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science*, *345*(6199), 484-501.

King, G., & Zeng, L. (2001). Logistic Regression in Rare Events Data, *Society for Political Methodology 7(*27), 137-163.

Maitra, S. (2021, September 29). State Space Model and Kalman Filter for Time-Series Prediction. *Medium*. Retrieved June 3, 2022, from https://towardsdatascience.com/state-space-model-and-kalman-filter-for-time-series-prediction-basic-structural-dynamic-linear-2421d7b49fa6

Matyasik, M. (2014). Secure sustainable development: Impact of social media on political and social crises. *Journal of Security and Sustainability Issues*, *4*(1), 5–16.

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic Polarization, Potential Conflict, and Civil Wars. *American Economic Review*, *95*(3), 796–816.

Munger, K. (2019). Social Media, Political Science, and Democracy. *The Journal of politics*, *81*(3), 62–67.

Olar, R.-G. (2019). Do they know something we don’t? Diffusion of repression in authoritarian regimes. *Journal of Peace Research*, *56*(5), 667–681.

Olson, M. (1989). Collective Action. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The Invisible Hand*, The New Palgrave (pp. 61–69). London: Palgrave Macmillan UK.

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, *96*(1), 3–14.

Pierskalla, J. H., & Hollenbach, F. M. (2013). Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa. *The American political science review*, *107*(2), 207–224.

Posen, B. R. (1993). The security dilemma and ethnic conflict. *Survival*, *35*(1), 27–47.

Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. *Intermountain Journal of sciences*, *5*(1–4), 23–27.

Salehyan, I., & Hendrix, C. S. (2014). Climate shocks and political violence. *Global Environmental Change*, *28*, 239–250.

Sambanis, N. (2001). Do Ethnic and Nonethnic Civil Wars Have the Same Causes?: A Theoretical and Empirical Inquiry (Part 1). *Journal of Conflict Resolution*, *45*(3), 259–282.

Sambanis, N., & Schulhofer-Wohl, J. (2019). Sovereignty Rupture as a Central Concept in Quantitative Measures of Civil War. *Journal of Conflict Resolution*, *63*(6), 1542–1578.

Steinert-Threlkeld, Z. C., Mocanu, D., Vespignani, A., & Fowler, J. (2015). Online social networks and offline protest. *EPJ data science*, *4*(1), 1–9.

Stockemer, D. (2010). Regime type and civil war – a re-evaluation of the inverted U-relationship. *Global Change, Peace & Security*, *22*(3), 261–274. Routledge.

Sunstein, C. (2018). Is Social Media Good or Bad for democracy? *International Journal of Human Rights*, *15*(27), 83–89.

Tähtinen, T. (2021). *When Facebook Is the Internet: The Role of Social Media in Ethnic Conflict* (No. EUI ECO 2021/01). *Economics Working Papers*, Economics Working Papers. European University Institute. Retrieved March 24, 2022, from https://ideas.repec.org/p/eui/euiwps/eco2021-01.html

Tang, S. (2015). The Onset of Ethnic War: A General Theory. *Sociological theory*, *33*(3), 256–279.

Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of communication*, *62*(2), 363–379.

Tunzelmann, A. von. (2022, April 25). The big idea: Can social media change the course of war? *The Guardian*. Retrieved May 13, 2022, from https://www.theguardian.com/books/2022/apr/25/the-big-idea-can-social-media-change-the-course-of-war

Vogt, G., Hunziker, P., Cederman, L.-E., Bormann, N.-C., Rüegger, S., & Vogt, M. (2021). GROWup—Geographical Research On War, Unified Platform. ETH Zurich. Retrieved from http://growup.ethz.ch/

Wegenast, T. C., & Basedau, M. (2014). Ethnic fractionalization, natural resources and armed conflict. *Conflict Management and Peace Science, 31*(4), 432–457.

Weidmann, N. B. (2009). Geography as Motivation and Opportunity: Group Concentration and Ethnic Conflict. *Journal of Conflict Resolution*, *53*(4), 526–543.

Weidmann, N. B. (2015). Communication networks and the transnational spread of ethnic conflict. *Journal of Peace Research*, *52*(3), 285–296.

Weidmann, N. B. (2016). A Closer Look at Reporting Bias in Conflict Event Data. *American Journal of Political Science*, *60*(1), 206–218.

Weidmann, N. B., & Rød, E. G. (2019). *The Internet and political protest in autocracies*. Oxford studies in digital politics. New York, NY: Oxford University Press.

World Bank. (2022a). Individuals using the Internet (% of population) | Data. *Individuals using the Internet*. Retrieved May 10, 2022, from https://data.worldbank.org/indicator/IT.NET.USER.ZS

World Bank. (2022b). World Bank Open Data | Data. Retrieved May 12, 2022, from https://data.worldbank.org/

Yardi, S., & Boyd, D. (2010). Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science, Technology & Society*, *30*(5), 316–327.

8. Appendix

8.1 Appendix 1: Imputation of Data

The imputation for social media data was done based on the existent data for the years 2017-2022. These six data points were used with Kalman smoothing in order to predict the previous values. "Kalman filter algorithm uses a series of measurements observed over time, containing noise and other inaccuracies, and produces estimates of unknown variables. This estimate tends to be more accurate than those based on a single measurement alone" (Maitra, 2021). To see how the Kalman smoothing was applied for some cases, figure A and figure B show the imputed data for the ten countries with the highest social media penetration and lowest social media penetration respectively. While error in the data is unavoidable, it is expected that the imputed data does differ from the real values to some extent. Especially, with such limited data and a large number of datapoints to be imputed, the imputed data might vary significantly from the real data. Comparing the average of the imputed data for 2010 and 2005 with the average of social media penetration worldwide (which is known), it becomes clear that the Kalman smoothing overestimates the amount of social media penetration. Additionally, while Kalman smoothing takes variation in the data into account, large and unpredictable changes to the data cannot be estimated. Additionally, for Kalman smoothing to be applied, a start values have to be chosen. For simplicity, this start value was 0% internet penetration in every country in the year 2000.

Figure A

Social Media Penetration imputed data in the countries with the lowest social media penetration

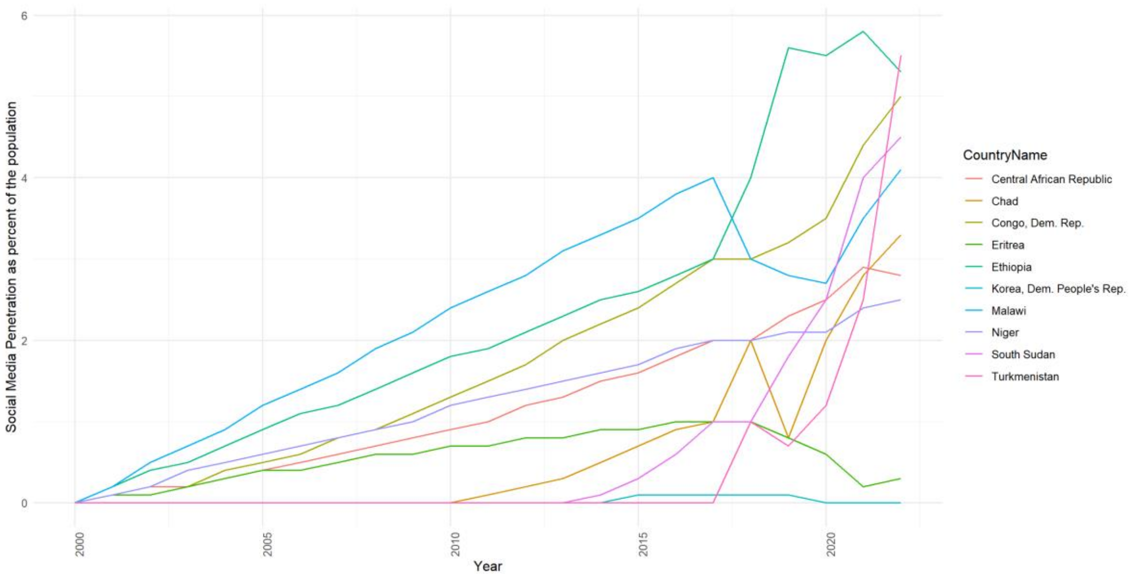
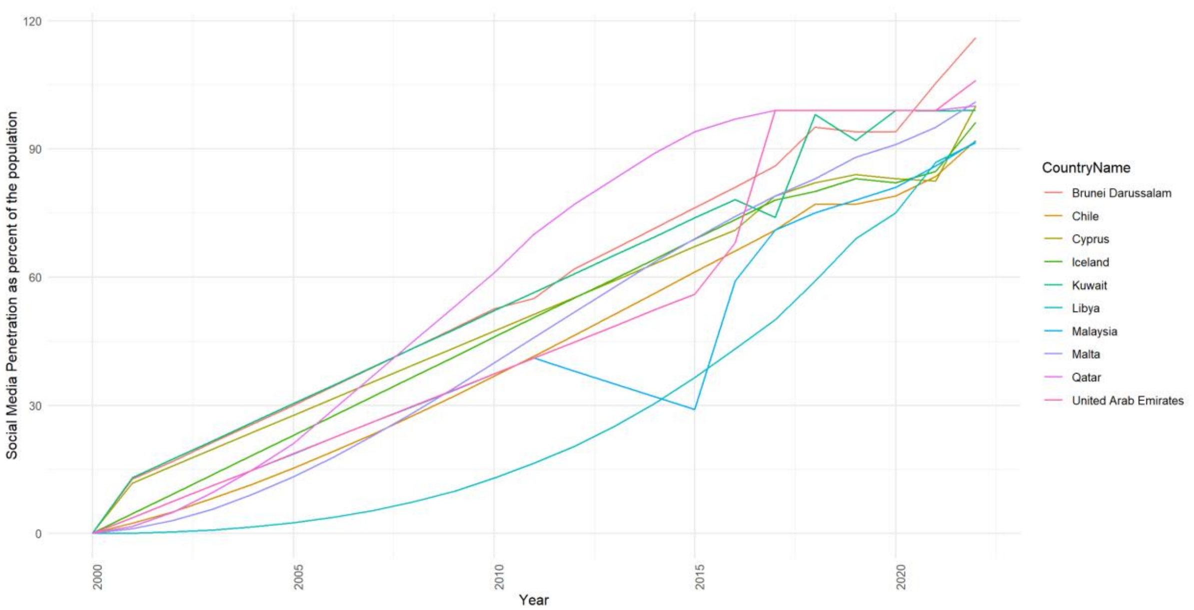


Figure B

Social Media Penetration imputed data in the countries with the highest social media penetration



8.2 Appendix 2: Table A: Other Models to Check Robustness of the Employed Models [Odds Ratio with 95%CI]

	All Log	Ethnic Log	Non-Ethnic Log	All 2005-2020	Ethnic 2005-2020	Non-ethnic 2005-2020	All 2000-2019	Ethnic 2000-2019	Non-ethnic 2000-2019
	[Model 9]	[Model 10]	[Model 11]	[Model 12]	[Model 13]	[Model 14]	[Model 15]	[Model 16]	[Model 17]
Internet Access	1.005	0.996	1.015	1.008	0.999	1.012	1.017*	1.008	1.031*
	[0.990, 1.020]	[0.976, 1.017]	[0.994, 1.036]	[0.990, 1.026]	[0.973, 1.025]	[0.989, 1.036]	[1.000, 1.034]	[0.984, 1.031]	[1.007, 1.055]
Regime Type	0.942	0.926	0.866	0.766	0.670	0.778	0.914	0.861	0.814
	[0.687, 1.290]	[0.598, 1.434]	[0.561, 1.338]	[0.544, 1.078]	[0.409, 1.096]	[0.499, 1.211]	[0.672, 1.244]	[0.566, 1.310]	[0.531, 1.249]
Log_GDP	0.613*	0.491*	0.700						
	[0.404, 0.930]	[0.283, 0.852]	[0.378, 1.298]						
GDP_per_Capita_PPP				1.000**	1.000*	1.000*	1.000***	1.000**	1.000*
				[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Military expenditure	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
GDP_Growth_Annual	1.029	1.041	0.964	0.993	0.985	0.971	1.030	1.030	0.987
	[0.980, 1.079]	[0.985, 1.101]	[0.891, 1.043]	[0.935, 1.054]	[0.914, 1.062]	[0.895, 1.054]	[0.982, 1.080]	[0.976, 1.086]	[0.906, 1.074]
Log_Population	1.891***	1.898***	1.888***						
	[1.598, 2.239]	[1.494, 2.411]	[1.501, 2.375]						
Population				1.000***	1.000**	1.000	1.000***	1.000***	1.000*
				[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rural Population	0.986	0.976	0.999	0.995	0.987	1.005	0.993	0.985	1.008
	[0.967, 1.006]	[0.950, 1.003]	[0.973, 1.025]	[0.979, 1.011]	[0.964, 1.012]	[0.984, 1.027]	[0.979, 1.008]	[0.965, 1.006]	[0.986, 1.030]
Rugged Terrain	1.164	1.371*	0.894	1.064	1.222	0.871	1.045	1.192	0.803
	[0.924, 1.465]	[1.004, 1.873]	[0.638, 1.255]	[0.846, 1.337]	[0.873, 1.710]	[0.640, 1.185]	[0.841, 1.300]	[0.891, 1.595]	[0.579, 1.114]
Pop Growth Rate	1.223*	1.028	1.453***	1.448**	1.221	1.602**	1.406**	1.269	1.529**
	[1.033, 1.447]	[0.748, 1.414]	[1.170, 1.805]	[1.157, 1.813]	[0.846, 1.764]	[1.208, 2.124]	[1.135, 1.740]	[0.928, 1.734]	[1.152, 2.030]
Natural Ressource Rent	1.017	1.043**	0.962	1.009	1.038**	0.951*	1.020	1.042***	0.967
	[0.995, 1.040]	[1.016, 1.072]	[0.922, 1.005]	[0.986, 1.033]	[1.012, 1.066]	[0.908, 0.997]	[0.999, 1.041]	[1.018, 1.067]	[0.925, 1.010]
Young Male Population	1.146*	0.998	1.322**	0.971	0.870	1.049	1.009	0.908	1.139
	[1.009, 1.302]	[0.848, 1.175]	[1.083, 1.614]	[0.821, 1.147]	[0.695, 1.088]	[0.833, 1.321]	[0.869, 1.170]	[0.752, 1.096]	[0.908, 1.429]
(Constant)	0.00000***	0.0002*	0.00000***	0.095	0.690	0.011	0.022*	0.141	0.001**
	[0.000, 0.001]	[0.00000, 0.341]	[0.000, 0.0001]	[0.003, 3.076]	[0.007, 67.958]	[0.0001, 1.380]	[0.001, 0.451]	[0.004, 5.671]	[0.00001, 0.109]
N	2,819	2,819	2,819	2,121	2,121	2,121	2,644	2,644	2,644

Table B: Logistic Regression including alternative form of information index (Odds Ratio with 95% CI)

	Onset All Conflict [Model 18]	Onset Ethnic Conflict [Model 19]	Onset Non Ethnic-Conflict [Model 20]
Alternative source of information	1.009	0.992	1.016
	[0.972, 1.047]	[0.965, 1.019]	[0.990, 1.042]
Regime Type	0.996	1.000	0.991*
	[0.985, 1.007]	[0.992, 1.008]	[0.984, 0.999]
GDP_per_Capita_PPP	1.000*	1.000**	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Military expenditure	1.000	1.000	1.000***
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
GDP_Growth_Annual	1.001	1.001*	1.000
	[1.000, 1.002]	[1.000, 1.002]	[0.999, 1.001]
Population	1.000***	1.000	1.000*
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rural Population	1.000	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rugged Terrain	1.001	1.003	0.999
	[0.995, 1.006]	[0.999, 1.007]	[0.996, 1.003]
Pop Growth Rate	1.007*	1.002	1.003
	[1.001, 1.012]	[0.999, 1.006]	[1.000, 1.007]
Natural Resourcese Rent	1.000	1.001***	0.999*
	[1.000, 1.001]	[1.000, 1.001]	[0.999, 1.000]
YoungMalePopulation15to24	1.000	0.998	1.001
	[0.997, 1.003]	[0.996, 1.000]	[0.999, 1.003]
(Constant)	1.020	1.050	0.997
	[0.949, 1.095]	[1.000, 1.104]	[0.950, 1.047]
N	3,035	2,643	2,711

Note: Odds Ratio with 95% confidence interval in brackets

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

8.3 Appendix 3: Models original data social media penetration

Table C: Logistic Regression Results: How does Social Media Penetration (Original Data 2017-2020) influences the Onset of Intrastate Conflict (Odds Ratio with 95%CI Interval)

	Intrastate conflict onset [Model 21]	Ethnic intrastate onset [Model 22]	Non-ethnic intrastate onset [Model 23]
Social Media Penetration Original (2017-2020)	0.988	1.021	0.963
	[0.956, 1.026]	[0.958, 1.076]	[0.929, 1.022]
Regime Type: [Ref. = Closed Autocracy]			
Electoral Autocracy	1.296	34,191,865.000	0.542
	[0.417, 61.354]	[0.000, Inf.000]	[0.139, 40.032]
Electoral Democracy	0.728	10,946,553.000	0.371
	[0.237, 34.574]	[0.000, Inf.000]	[0.095, 32.761]
Liberal Democracy	1.540	71,589,323.000	0.000
	[0.065, 39.238]	[0.000, Inf.000]	[0.000, Inf.000]
GDP per Capita PPP	1.000	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Military expenditure	1.000	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
GDP Growth Annual	1.090	1.110	1.095
	[0.994, 1.192]	[0.943, 1.310]	[0.968, 1.212]
Population	1.000	1.000	1.000
	[1.000, 1.000]	[1.000, 1.000]	[1.000, 1.000]
Rural Population	0.988	0.966	0.997
	[0.949, 1.018]	[0.896, 1.025]	[0.951, 1.034]
Rugged Terraine	0.940	1.449	0.698
	[0.671, 1.570]	[0.819, 3.264]	[0.473, 1.457]
Pop-Growth Rate	1.754	2.779	1.339
	[1.014, 3.383]	[0.794, 10.081]	[0.772, 3.041]
Natural Resources Rent	1.018	1.036	1.005
	[0.953, 1.047]	[0.954, 1.102]	[0.921, 1.046]
YoungMalePopulation15to24	1.161	1.169	1.230
	[0.770, 1.550]	[0.622, 2.519]	[0.691, 1.629]
(Constant)	0.001	0.000	0.002
	[0.00000, 4.391]	[0.000, Inf.000]	[0.00000, 149.536]
N	567	567	567

Note: Odds Ratio with 95% confidence interval in brackets

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

8.4 Appendix 4: R-Code Regression and Data preparation

```
#### Libraries ####
library(haven)
library(car)
library(psych)
library(arm)
library(ggplot2)
library(sjPlot)
library(olsrr)
library(dplyr)
library(visreg)
library(effectsize)
library(ggpubr)
library(ggdist)
library(tidyverse)
library(broom)

theme_set(theme_sjplot())

#### Dataset ####
Master_Thesis_Dataset_with_Control_Variables <- read_sav("Master Thesis Dataset with Control Variables.sav")
View(Master_Thesis_Dataset_with_Control_Variables)

dataset <- Master_Thesis_Dataset_with_Control_Variables
attach(dataset)

View(dataset)

attach(data.frame(dataset))

#### Filter Old Cases####
```

```
dataset = filter(dataset, Year > 1999)


#### Defining Categorical Variables ####
dataset$v2x_regime <- factor(dataset$v2x_regime, levels = c(0,1,2,3),labels = c("Closed Autocracy", "Elecotral Autocracy", "Electoral Democracy", "Liberal Democracy"))
is.factor(dataset$v2x_regime)
levels


dataset$onset_ko_flag <- factor(dataset$onset_ko_flag)
is.factor(dataset$onset_ko_flag)


dataset$onset_ko_eth_flag <- factor(dataset$onset_ko_eth_flag)
is.factor(dataset$onset_ko_flag)


dataset$onset_ko_noneth_flag <- factor(dataset$onset_ko_noneth_flag)
is.factor(dataset$onset_ko_flag)


dataset$plot <- factor(dataset$plot)
is.factor(dataset$plot)



####Descriptive Statistics####
attach(dataset)
summary(dataset)
describe(dataset)


table(dataset$onset_ko_flag)
table(dataset$onset_ko_eth_flag,dataset$Year)
table(dataset$onset_ko_noneth_flag,dataset$Year)
table(dataset$v2x_regime,dataset$Year)
table(dataset$plot,dataset$Year)
```

```
####Descriptive Plots####
#Internet Access over the years#
ggplot(dataset, aes(x = factor(Year), y = Internet_Access)) +
  geom_bar(stat = "summary", fun = "mean", fill = "steelblue")+
  theme_set(theme_minimal())+
  labs(x = "Year", y = "Internet Access in percent of the global population")

ggplot(data = na.omit(dataset), aes(x = factor(v2x_regime), y = Internet_Access)) +
  geom_bar(stat = "summary", fun = "mean", fill = "steelblue")+
  theme_set(theme_minimal())+
  labs(x = "Regime Type", y = "Internet Access in %")

#Different Conflict over the years#

ggplot(data=subset(dataset,!is.na(onset_ko_flag)), aes(x = Year, fill = onset_ko_flag)) +
  geom_bar(stat = "count") +
  scale_fill_manual(values = c("white", "steelblue", "grey"),labels = c("", "Conflict Onset", "Missing Data"))+
  scale_x_continuous(breaks = seq(2000,2020))+
  labs(x="Year", y= "Number of Conflict Onsets")+
  ylim(0,20)+
  theme_set(theme_minimal())+
  guides(fill=guide_legend(title=NULL))

ggplot(data=subset(dataset,!is.na(plot)), aes(x = Year, fill = plot)) +
  geom_bar(stat = "count") +
  scale_x_continuous(breaks = seq(2000,2020))+
  labs(x="Year", y= "Number of Conflict Onsets by Ethnic and Non-Ethnic Conflict")+
  ylim(0,20)+
  scale_fill_manual(values = c("white", "steelblue", "lightblue"),labels = c("", "Non-Ethnic Conflict", "Ethnic Conflict"))+
  theme_set(theme_minimal())+
```

```
  guides(fill=guide_legend(title=NULL))



####Assumption Checks####

#Multicolineraity#

vif(M1)

vif(M2)

vif(M3)

vif(M4)


#### Logistic Regression Model####


library(jtools)

library(stargazer,type = "text")

M1 <- glm(onset_ko_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24, family=binomial, data=dataset )

summary(M1)

round(M1$coefficients, digits = 3)

round(M1$std.error)


M2 <- glm(onset_ko_eth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24, family=binomial, data=dataset )

summary(M2)

summ(M2)

tidy(M2)


M3 <- glm(onset_ko_noneth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24, family=binomial, data=dataset )

summary(M3)

tidy(M3)
```

```
####Table####

library(stargazer)

#Coefficents#

stargazer(M1,M2,M3,M4, type = "text", keep.stat = "n", style = "apsr", star.cutoffs = c(0.05, 0.01, 0.001), out="C://Users//baier//OneDrive//Desktop//Name.html")


#Odds Ratio and Confidence Intervalls#

CI.OR1 <- as.matrix(exp(confint.default(M1)))

CI.OR2 <- as.matrix(exp(confint.default(M2)))

CI.OR3 <- as.matrix(exp(confint.default(M3)))

CI.OR4 <- as.matrix(exp(confint.default(M4)))

stargazer(M1,M2,M3,M4, apply.coef = exp, ci.custom = list(CI.OR1,CI.OR2,CI.OR3,CI.OR4), type = "text", keep.stat = "n", style = "apsr", star.cutoffs = c(0.05, 0.01, 0.001), out="C://Users//baier//OneDrive//Desktop//Name.html", t.auto=F, p.auto=F, ci = T)



#### Odds Ratio All Models ####

OddsM1 <- exp(coefficients(M1))

round(OddsM1, digits = 3)


OddsM2 <- exp(coefficients(M2))

round(OddsM2, digits = 3)


OddsM3 <- exp(coefficients(M3))

round(OddsM3, digits = 3)


#### Plot of the regression ####


plot_model(M1, grid = TRUE)+labs(title = "Odds Ratio of Model 1")


plot_model(M2)
```

```
plot_model(M3)



visreg(M1, "Internet_Access",
       gg = TRUE,
       scale="response") +
  labs(y = "Probability (Onset of Conflict) ",
       x = "Internet Access in percent of the whole population",
       title = "Relationship between Internet Access and Onset of Conflict",
       subtitle = "including all control variables",
       caption = "")

visreg(M2, "Internet_Access",
       gg = TRUE,
       scale="response") +
  labs(y = "Probability (Onset of Ethnic Conflict)",
       x = "Internet Access in percent of the whole population",
       title = "Relationship between Internet Access and Onset of Ethnic Conflict",
       subtitle = "including all control variables",
       caption = "")



visreg(M3, "Internet_Access",
       gg = TRUE,
       scale="response") +
  labs(y = "Probability (Onset of Non-Ethnic Conflict)",
       x = "Internet Access in percent of the whole population",
       title = "Relationship between Internet Access and Onset of Non-Ethnic Conflict",
       subtitle = "including all control variables",
       caption = "")
```

```
#### Making the plots look nice####
 p1 <- visreg(M1, "Internet_Access",
             gg = TRUE,
             scale="response",
             partial = TRUE,
             rug = TRUE,
             plot = FALSE)

p2 <- visreg(M2, "Internet_Access",
               gg = TRUE,
               scale="response",
               partial = TRUE,
               rug = TRUE,
               plot = FALSE)

p3 <- visreg(M3, "Internet_Access",
             gg = TRUE,
             scale="response",
             partial = TRUE,
             rug = TRUE,
             plot = FALSE)



####Final Plots###
#Plot Model 1#
ggplot(p1$fit, aes(Internet_Access, visregFit)) +
  geom_ribbon(aes(ymin=visregLwr, ymax=visregUpr),  alpha=0.05,
              colour="orange", linetype=1, size=0.1) +
  geom_line(colour = "red") +
  stat_dots(data = dataset,aes(y = onset_do_flag, side = ifelse(onset_do_flag == 0, "top", "bottom")),scale = 0.7)+
  labs(y = "P(Onset of Conflict)", x = "İnternet Access in Percent of the population")+
  scale_y_continuous(name= "P(Onset of Conflict)")+
```

```
  labs(title = "Logistic Regression Model: Internet Access ~ Onset of Conflict" )


#Plot Model 2#
ggplot(p2$fit, aes(Internet_Access, visregFit)) +
  geom_ribbon(aes(ymin=visregLwr, ymax=visregUpr),  alpha=0.2,
              colour="lightblue", linetype=1, size=0.2) +
  geom_line(colour = "red") +
  stat_dots(data = dataset,aes(y = onset_do_flag, side = ifelse(onset_do_flag == 0, "top", "bottom")),scale = 0.7)+
  labs(y = "P(Onset of Conflict)", x = "Ïnternet Access in Percent of the population")+
  scale_y_continuous(name= "P(Onset of Ethnic Conflict)")+
  labs(title = "Logistic Regression Model: Internet Access ~ Onset of Ethnic Conflict" )



#Plot Model 3#
ggplot(p3$fit, aes(Internet_Access, visregFit)) +
  geom_ribbon(aes(ymin=visregLwr, ymax=visregUpr),  alpha=0.2,
              colour="lightblue", linetype=1, size=0.2) +
  geom_line(colour = "red") +
  stat_dots(data = dataset,aes(y = onset_do_flag, side = ifelse(onset_do_flag == 0, "top", "bottom")),scale = 0.7)+
  labs(y = "P(Onset of Non-Ethnic Conflict)", x = "Ïnternet Access in Percent of the population")+
  scale_y_continuous(name= "P(Onset of Non- Ethnic Conflict)")



#### Comparison Ethnic Conflict and Non-Ethnic Conflict####
#With Confidence Interval#
ggplot(p2$fit, aes(Internet_Access, visregFit)) +
  geom_ribbon(p2$fit, mapping = aes(ymin=visregLwr, ymax=visregUpr),  alpha=0.05,
              colour="orange", linetype=2, size=0.2) +
  geom_line(data=p2$fit, aes (colour = "Ethnic Conflict"))+
  geom_ribbon(p3$fit, mapping = aes(ymin=visregLwr, ymax=visregUpr), alpha = 0.05, color = "lightblue", linetype = 2, size = 0.2)+
```

```
  geom_line(data=p3$fit, aes (colour ="Non-Ethnic Conflict"))+
  scale_colour_manual("Legend", values = c("red","blue"))+
  labs(y = "P(Onset of (Ethnic) Conflict", x = "Internet Access in Percent of the Population")


#Without confidence Interval#
ggplot(p2$fit, aes(Internet_Access, visregFit)) +
  geom_line (aes(color = "Ethnic Conflict"))+
  geom_line(data=p3$fit, aes( color ="Non-Ethnic Conflict"))+
  labs(y = "P(Onset of (Ethnic) Conflict", x = "Internet Access in Percent of the Population")+
  scale_colour_manual("Legend", values = c("red","blue"))



####Model Fit####
#Deviance of the Models (Measure of Error)#
library(lmtest)

logistic_model1 <- M1
logistic_model1$deviance

logistic_model2 <- M2
logistic_model2$deviance

logistic_model3 <- M3
logistic_model3$deviance

#Likelihood Ratio Test (comparing if adding predictors increases the explanatory power)#
lrtest(M1,M2,M3)
lrtest(M1)
logLik(M4)
```

```
#Accuracy#
library(performance)
performance_pcp(M1, ci = 0.95, method = "herron", verbose = TRUE)
performance_pcp(M2, ci = 0.95, method = "herron", verbose = TRUE)
performance_pcp(M3, ci = 0.95, method = "herron", verbose = TRUE)
performance_pcp(M4, ci = 0.95, method = "herron", verbose = TRUE)

#Pseudo R² (between 0.2 - 0.4 "excellent fit")#
library(DescTools)
PseudoR2(M1,which="Nagelkerke")
PseudoR2(M2,which="Nagelkerke")
PseudoR2(M3,which="Nagelkerke")
PseudoR2(M4,which="Nagelkerke")

PseudoR2(M1,which="CoxSnell")
PseudoR2(M2,which="CoxSnell")
PseudoR2(M3,which="CoxSnell")
PseudoR2(M4,which="CoxSnell")



library(pscl)
pR2(M1)
pR2(M2)
pR2(M3)
pR2(M4)

#Wald Test (statistical significant of each coefficient)#
library(survey)
regTermTest(M1, "Internet_Access")

regTermTest(M3, "Internet_Access")
```

```
#Importance of Variables (absolute t values)#
library(caret)
varImp(M1)

#Plotting the importance of variables#
V = caret::varImp(M3)

ggplot2::ggplot(V, aes(x=reorder(rownames(V),Overall), y=Overall)) +
  geom_point( color="blue", size=4, alpha=0.6)+
  geom_segment( aes(x=rownames(V), xend=rownames(V), y=0, yend=Overall),
              color='skyblue') +
  xlab('Variable')+
  ylab('Overall Importance')+
  theme_light() +
  coord_flip()



####Interaction Effect####

#Model with Interaction effect#
M4 <- glm(onset_ko_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 + Internet_Access*EFindex, family=binomial, data=dataset )
M5 <- glm(onset_ko_eth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 + Internet_Access*EFindex, family=binomial, data=dataset )
M6 <- glm(onset_ko_noneth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 + Internet_Access*EFindex, family=binomial, data=dataset )

summary(M4)
summary(M5)
summary(M6)
```

```
M7<- glm(onset_ko_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 + EFindex, family=binomial, data=dataset )

M8 <- glm(onset_ko_eth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 + EFindex, family=binomial, data=dataset )

M9 <- glm(onset_ko_noneth_flag ~ Internet_Access + v2x_regime + GDP_per_Capita_PPP + Military_expenditure + GDP_Growth_Annual +Population+RuralPopulation+Rugged_Terraine+PopGrowthRate+NaturalResourceseRent+ YoungMalePopulation15to24 +EFindex, family=binomial, data=dataset )


summary(M7)
summary(M8)
summary(M9)


#Odds Ratio Model 4#
OddsM4 <- exp(M4$coefficients)
OddsM4
OddsM5 <- exp(M5$coefficients)
OddsM5
OddsM6 <- exp(M5$coefficients)
OddsM6
stargazer(M4,M5,M6)
#Plot Interaction Effect#


#M4#
set_theme(theme_minimal())
Plot_Interaction <- plot_model (M4, type = "int")
Plot_Interaction +
  ylim(0,0.025)+
  labs(x = "Internet Access in Percent", y = "P(Onset of Conflict)", title = "")


#M5#
```

```
set_theme(theme_minimal())

Plot_Interaction <- plot_model (M5, type = "int")

Plot_Interaction +

  labs(x = "Internet Access in Percent", y = "P(Onset of Conflict)")+

  labs(title = "Predicted probabilities for the Onset of Ethnic Intrastate Conflict", subtitle = "by Ethnic Fracternalization")


#M6#

set_theme(theme_minimal())

Plot_Interaction <- plot_model (M6, type = "int")

Plot_Interaction +

  ylim(0,0.035)+

  labs(x = "Internet Access in Percent", y = "P(Onset of Conflict)")+

  labs(title = "Predicted probabilities for the Onset of Non-Ethnic Intrastate Conflict", subtitle = "by Ethnic Fracternalization")


####Log of big values####

dataset <- dataset %>%

  mutate(Log_GDP = log(GDP_per_Capita_PPP))


dataset <- dataset %>%

  mutate(Log_Population = log(Population))


####Missing Data####

view(dataset)

dataset.test <- subset(dataset, select = c(onset_ko_flag,Internet_Access, v2x_regime,GDP_per_Capita_PPP, Military_expenditure, GDP_Growth_Annual, Population, RuralPopulation, Rugged_Terraine, PopGrowthRate, NaturalResourceseRent, YoungMalePopulation15to24 ))

view(dataset.test)

dataset.test <- (na.omit(dataset.test))

table(dataset.test$onset_ko_flag)
```

8.5 Appendix 5: R Code data imputation

```
#### Interpolating Missing Data - Social Media ####

#Import Social Media Data#

library(readxl)


#Social_Media_Access_Own_Dataset <- read_excel("C:/Users/baier/OneDrive/Desktop/Social Media Access Own Dataset.xlsx")

View(Social_Media_Access_Own_Dataset)

countries <- read_excel("C:/Users/baier/OneDrive/Desktop/countries.xlsx")


df1 <- countries


x1 <-  as.numeric(df1)


library(imputeTS)


df.impute <- na_kalman(x1)


data.frame(df.impute)


df.impute <- t(df.impute)


View(df.impute)



warnings()


#### Doing it row by row####

Social_Media_Access_Own_Dataset <- read_excel("C:/Users/baier/OneDrive/Desktop/Social Media Access Own Dataset.xlsx")

df <- Social_Media_Access_Own_Dataset

View(df)
```

```
v1 <- as.numeric(df[173,2:24])
impute173 <- na_kalman(v1)
class.df <- cbind(impute1, impute2, impute3,impute4, impute5, impute6, impute7,
          impute8, impute9, impute10, impute11, impute12, impute13, impute14,
          impute15, impute16, impute17, impute18, impute19, impute20, impute21,
          impute22, impute23, impute24, impute25, impute26, impute27,
          impute28, impute29, impute30, impute31, impute32, impute33,
          impute34, impute35, impute36, impute37, impute38, impute39,
          impute40, impute41, impute42, impute43, impute44, impute45,
          impute46, impute47, impute48,impute49, impute50, impute51,
          impute52,impute53, impute54,impute55, impute56, impute57,
          impute58, impute59, impute60, impute61, impute62, impute63,
          impute64, impute65, impute66, impute67, impute68, impute69,
          impute70, impute71, impute72, impute73, impute74, impute75,
          impute76, impute77,impute78, impute79, impute80, impute81, impute82,
          impute83, impute84,impute85, impute86, impute87, impute88,
          impute89, impute90, impute91, impute92, impute93, impute94,
          impute95, impute96, impute97, impute98, impute99, impute100,
          impute101, impute102, impute103, impute104, impute105, impute106, impute107,
          impute108, impute109, impute110, impute111, impute112, impute113, impute114,
          impute115, impute116, impute117, impute118, impute119, impute120, impute121,
          impute122, impute123, impute124, impute125, impute126,impute127, impute128,
          impute129, impute130, impute130,impute132, impute133, impute134, impute135,
          impute136, impute137, impute138, impute139, impute140, impute141, impute142,
          impute143, impute144, impute145, impute146, impute147, impute148, impute149,
          impute150, impute151, impute152, impute153, impute154, impute155, impute156,
          impute157, impute158, impute159, impute160, impute161, impute162, impute163,
          impute164, impute165, impute166,impute167, impute168, impute169, impute170,
          impute171, impute172, impute173)
```

```
#impute1 <- t(impute1)

View(class.df)

write.csv(class.df,"C:/Users/baier/OneDrive/Desktop/Imputed Data.csv", row.names = FALSE)


#### Doing it row by row####

library(dplyr)

df.new <- class.df


x <- ts(c(0,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,0.99,0.99,0.99,0.99,0.99,1))

na_kalman(x)

v1 <- as.numeric(df[56,2:24])

new56 <- na_kalman(v1)

df.new <- cbind(class.df, new56)

View(df.new)
```