Armed Conflict Location Event

Akira Sasaki, Eduardo Acostamadiedo, Oliver Recep Çok

11/24/2017

Data Cleaning

Since we wanted to find a way to measure the correlation of civilian based conflicts and Facebook penetration, we used the Armed Conflict Location & Event Data Project (ACLED) dataset set to determine which variables would be relevant. We needed to determine a time period to properly compare these conflicts with the Facebook user data. For this reason, we agreed to only select the "YEAR", "EVENT_TYPE", "COUNTRY" variables out of the ACLED dataset for our project. The "COUNTRY" variable became our limiting factor as most the countries in the world do not experience armed conflict.

To make the variables more accessible, we transformed both "YEAR" and "COUNTRY" into "year" and "country" respectively. The most important variable that we had to transform was the "EVENT_TYPE". We had to determine which armed conflict events would be classified as either a "CIVILIAN INVOLVED CONFLICT" or an "ARMED GROUP CONFLICT". After determining this we then mutated these different types into either one of these two classifications, which would become our dependent variables. We implemented all of these changes into a new dataset called ACLED2.

Since there were still around 4000 observations of different instances of armed conflicts in ACLED2, we wanted to compile the dataset by year and country. We tied all the armed conflicts in each country for that year into one row. This significantly decreased the number of rows in ACLED2, which increased the readability of the data.

In the GDP dataset we needed to have a variable that could be consistent enough to act as a control, which is why we needed to create a GDP per capita variable for each country overtime. The GDP dataset had countries that were spelled differently from the ACLED2 dataset due to having phrases such as "Republic of" added. We recoded each of these country names to match the ACLED2 country names. Also, we recoded some of the variables in the GDP dataset that had capitalization differences from the variables in ACLED2. This enabled us to merge both the ACLED2 dataset and the GDP dataset into the new DATA dataset.

Within the new DATA dataset, we added the 3 letter country codes for each country as a new variable called "iso3c". This was to further simply searching for countries within the dataset. In addition to the already existing variables "CIVILIAN INVOLVED CONFLICT" and "ARMED GROUP CONFLICT", we wanted to create new variables that represented the rate of conflicts for these countries. By taking the natural log of the percentage of both previous variables, we created the "lnsocial" (for "CIVILIAN INVOLVED CONFLICT") and the "lnmilitary" (for "ARMED GROUP CONFLICT") variables for the DATA dataset.

Finally, we needed to add our social media variable to measure the relevance of social media to armed conflicts. Due to not finding the enough data for Twitter, we opted to collect Facebook data instead. By using cascading style sheets, we were able to web scrape data from web.archive.org to for Facebook users across countries. The data scraped was messy and required us to organize the data into a dataset of the number of Facebook users by country code overtime. To measure the relevance of the number of Facebook users (as the variable "subscriptions") for each country, we created a new variable called "penetration" to determine the number of Facebook users as a percentage of the population. After recoding a few country names in this dataset, we merged it by year and country code with the DATA dataset.

Explanatory Statitical Analysis

In the next section, we explore the relationship between social media and reporting of conflict events in two levels of analysis. The first level is a time comparison between 2010 and 2015. This contrast considers that the number of facebook users increased almost 192 percent in this short time. The second level is the type of conflict divided whether the conflict was civil society related or armed actor related. Between 2010 and 2015, civil associated conflicts increased 225 percent and armed actors 291 percent.

Although the reporting of both types of conflict increased in similar ways, as it will be shown, it seems that in contrast to armed actors conflict events, civil society events are more associated to the increase of Facebook users.

The 2015 civil society conflict events graph has an steeper slope than the 2010 plot, indicating that there is more association between facebook users and reports.

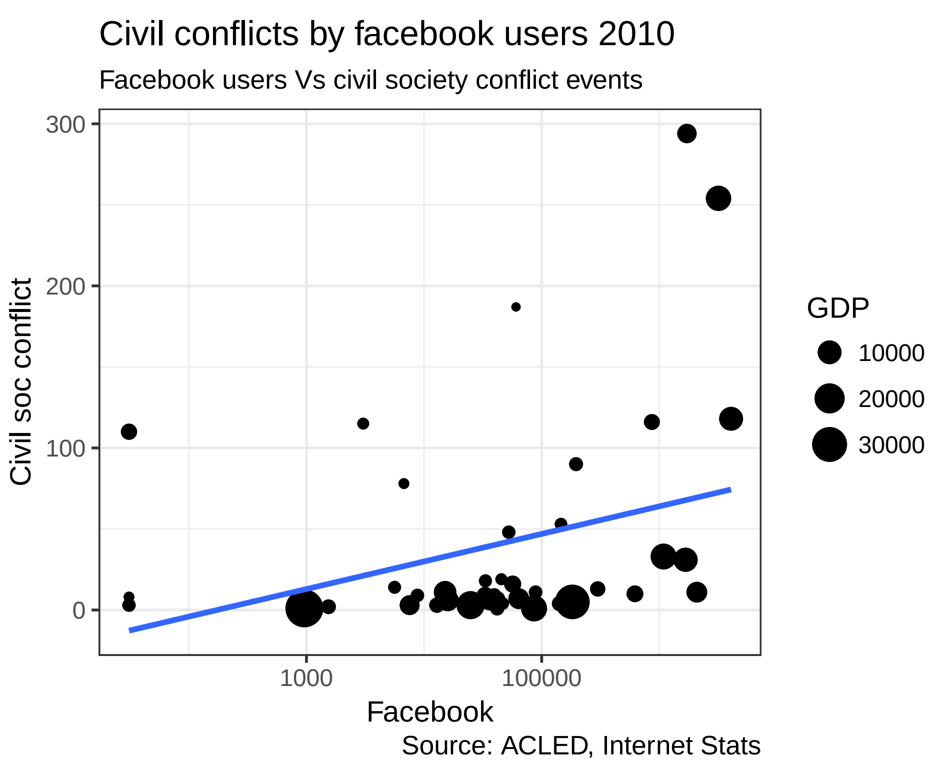
Turning to the armed actors conflicts, the 2010 and 2015 graph, the steepness of the line is almost the same, it seems that there was no significant association between facebook users and resporting of armed actors events.

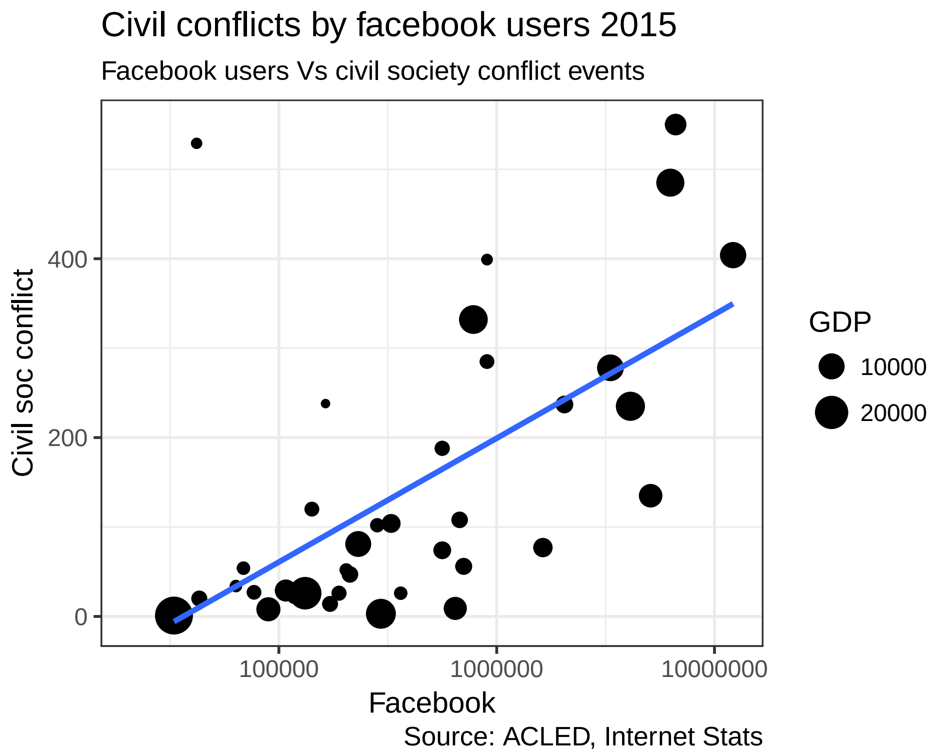
Table 1: Descriptive statistics 2010

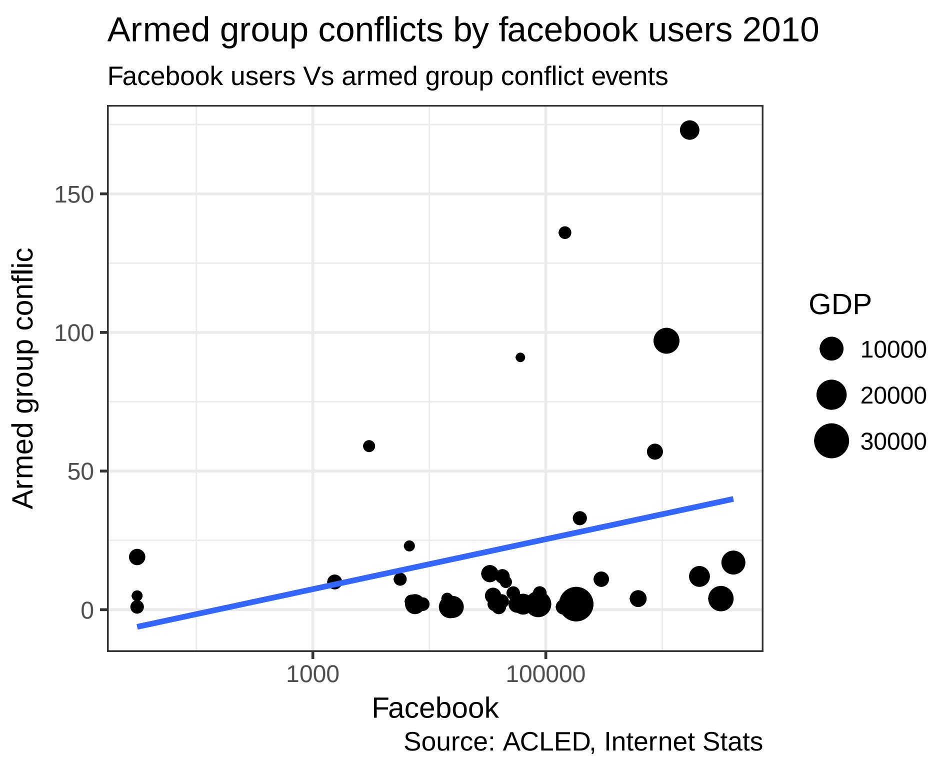
Statistic	N	Mean	St. Dev.	Min	Max
Armed actors events	47	22.9	41.9	0	173
Civilian events	47	43.1	67.6	0	294
GDP per cap.	47	5,075.4	7,187.9	596	37,141

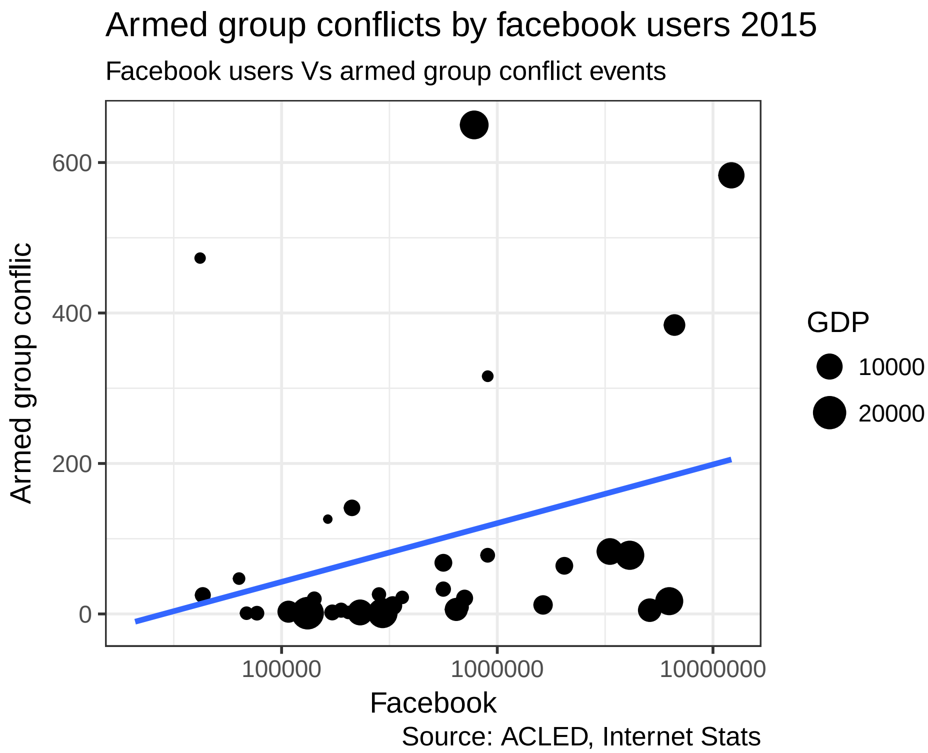
Table 2: Descriptive statistics 2015

Statistic	N	Mean	St. Dev.	Min	Max
Armed actors events	48	87.9	166.6	0	650
Civilian events	48	137.1	155.5	0	550
GDP per cap.	48	4,592.6	5,764.3	0	28,272




Civil conflicts by facebook users 2015
Facebook users Vs civil society conflict events
Civil soc conflict
400
200
0
100000
1000000
10000000
Facebook
Source: ACLED, Internet Stats
GDP
10000
20000


Armed group conflicts by facebook users 2010
Facebook users Vs armed group conflict events
Armed group conflic
150
100
50
0
1000
100000
Facebook
Source: ACLED, Internet Stats
GDP
10000
20000
30000



Regression Analysis

In our hypothesis, the prevalence of facebook should increase the number of reported conflict cases. In particular, between armed group conflicts and civilian conflicts (riots/protests and violance against civilians), we expect a strong correlation in civilian conflicts due to the nature of the conflicts. For civilian conflicts, it is important that people's voices are heard and attract attention.

At first, we examine the effect of facebook penetration on the number of reported conflict cases by OLS regression. We divide the conflicts into two different groups: civilian conflicts and armed group conflicts. Then, we applied the regression respectively. Also, for the sake of comparison, we decided to analyze the data from two different years: 2010 and 2015. The dependent variable is the number of reported conflict cases and our explanatory variable is the penetration of facebook in respective African countries. As a control variable, we included the natural logarith of GDP per capita as we consider the wealthiness at an individual level correlates with the penetration rate of facebook as well as the frequency of conflicts. The table is the results of respective OLS regression.

The first table shows the results of the OLS regression on the civilian conflict group.The coefficient of 2015 is bigger than that of 2010, meaning the facebook penetration in 2015 has a stronger magnitude on the increase of civilian conflict cases. However, both coefficients are not stastically significant hence they are unreliable. The impact of GDP per capita on the number of civilian conflicts are significantly different bewteen 2010 and 2015. In 2010 the coefficient was negative but statistically insignificant but in 2015 the coefficient became statistically very significant. According to the coefficient in 2015, 1% increase of GDP per capita increase the number of reported civilian conflicts by 13, which is a substantial magnitude.

The second table describes the results for the armed group conflicts. The coefficients of facebook penetration is in contrast to the previous group. In 2010 the coefficient was negative but in 2015 that turns into positive. The trend is in line with our assumption but why it was negative in 2010 is baffling. One possible explanation is in 2010 the people who had access to facebook could be part of armed group and they tried to conceal the conflicts. Nontheless, both coefficients are statistically insignificant and unreliable. GDP per capita is unsubstantial and statistically insignificant for both years.

Table 3: Civilan Conflicts

	Dependent variable:	
	2010	2015
	(1)	(2)
Facebook penetration	2.773	3.982
	p = 0.485	p = 0.210
GDP per cap. (log)	−3.565	13.071***
	p = 0.714	p = 0.007
Constant	67.145	25.739
	p = 0.369	p = 0.381
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 4: Armed Group Conflicts

	Dependent variable:	
	2010	2015
	(1)	(2)
Facebook penetration	−1.612	2.597
	p = 0.223	p = 0.617
GDP per cap. (log)	−1.542	9.931**
	p = 0.802	p = 0.024
Constant	37.611	5.138
	p = 0.420	p = 0.858
Note:	*p<0.1; **p<0.05; ***p<0.01	

We admit those regressions are susceptible to a omitted variable bias. We only included one control variable in the model but since we include many different countries, there so many factors which can play in. In order to control those different factors, we decided to do two-way fixed effect regression between countries and years. The result is below.

Table 5: Two-way Fixed Effect Regression

	Conflict type	
	Armed Group Conflict	Civilian Conflict
	(1)	(2)
Facebook penetration	0.095***	0.061*
	p = 0.008	p = 0.057
GDP per cap. (log)	0.017	0.001
	p = 0.752	p = 0.988
Observations	285	285
R^2	0.032	0.016
Adjusted R^2	−0.196	−0.215
F Statistic (df = 2; 230)	3.743**	1.847
Note:		*p<0.1; **p<0.05; ***p<0.01

Now it is possible to see interesting results. Facebook penetration became much more statistically signifant than the former models. Although, the magnitude of the coefficients are small, the positive correlation is in line with our assumption. However, the interesting finding is that the coefficient is more statistically significant for armed group conflict than civilian conflicts. This is opposite to our intuition. In contrast to facebook penetration, GDP per capita became immensely statistically insignificant and the magnitudes are very small as well.

Also, we found an issue with our model as well. The adjusted-R square became negative in this regression, indicating possibly a misspecification in the model. Also, F-statistics are small for both groups, meaning the model has a space of improvement. In the follwing, we discuss the possible correction in our model.

Conclusion and Future Study

In our