Political strongholds and the severity of organized violence during the 2008 Kenyan crisis*

Francisco Villamil†
ETH Zürich

November 15, 2014

Abstract

In this article I propose a new mechanism as part of the explanations for the violence dynamics during the 2008 Kenyan crisis. The basic working assumption is that the logic of organized violence is different from the violence in riots and protests and that the factors determining the incidence of that type of violence cannot account for its severity. Using data from the Armed Conflict & Location Data (ACLED) project, I show that the severity of violence was positively related to the degree of ethnic heterogeneity, but only in those areas where one of the parties had mass support. Results not only point to one of the multiple mechanisms explaining violence in Kenya, but also shed light on the logic of the local dynamics of communal violence and highlight the need for a complex disaggregation of violence measures according to the different actors involved in the process.

*Replication files at github.com/franvillamil/ov-kenya
†Contact at francisco.villamil@icr.gess.ethz.ch

Introduction

Kenya has been traditionally seen as an oasis of stability in East Africa. Compared to its direct neighbors, it has been free of major civil conflicts and has followed a steady path of economic development. However, during the first months of 2008 the country lived its worst political crisis since Independence. In the aftermath of the 27 December 2007 elections, a wave of popular protests emerged and promptly turned into an ethnic conflict, bringing about more than 1 000 deaths and around 300 000 displaced people in less than three months (Human Rights Watch, 2008).

The conflict originated in the context of a very tight electoral competition between challenger Raila Odinga (Orange Democratic Movement, ODM) and incumbent Mwai Kibaki (Party of National Unity, PNU). After a relatively peaceful campaign –at least compared to other episodes of pre-electoral violence in Kenya–, the first protests began when the announcement of results was delayed. Both the polls and the public opinion were pointing to Odinga as the winner, but the electoral commission announced on 30 December that Kibaki had won the election. To make things worse, the existence of widespread fraud favoring the incumbent candidate was widely known. In only a short time, the demonstrations turned into ferocious violence, and during the following weeks several parts of Kenya witnessed a wave of selective killings, directed at politically opposite ethnic groups. Violence continued until a power-sharing agreement was signed on 28 February between the PNU and the ODM (see e.g. Cheeseman, 2008; Mueller, 2008, for a general account of the conflict).

At first, the media viewed the conflict as the product of ancient hatreds, but this explanation has been regarded as excessively simplistic. On the contrary, violence has been related to several underlying factors. For instance, Mueller (2008) says that the electoral fraud was indeed the spark that triggered a conflict caused mainly by long-term dynamics, like grievances over land distribution, the deterioration of public institutions and the loss of state authority to local militias, and De Smedt (2009) focuses on the politicization of ethnicity and the 'big man' politics that came about since the start of multi-party democracy. Other approach regards violence as the consequence of Odinga's electoral campaign, focused on exploiting the tensions between the Kikuyu and other ethnic groups (Collier, 2009).

One of the riddles that come up when looking at the Kenyan crisis is the variation of violence levels throughout the country. Why were some district extremely violent while others remained relatively calm? Following the trend in conflict research of focusing on within-country dynamics (Kalyvas, 2008), this article tries to use this variation to test one of the possible mechanisms that explain the differential levels of violence.

The main argument is that the success of the recruitment process by the ethnic militias was higher on those areas strongly dominated by one of the two main political parties, and this in turn influenced the *severity* of organized violence across Kenya. To test this argument I use event data to be able to disaggregate violence not only in a spatial sense, but also in a ontological way. On one hand, I distinguish the agency of the violence, by taking into account only those events in which at least one of the sides was an organized armed group. On the other hand, I differentiate between the severity and incidence of violence, assuming that the factors behind its occurrence and its intensity are different. By proving this mechanism empirically, I contribute to both the particular explanations of the Kenyan crisis and the general understanding of collective action dynamics during ethnic conflict.

Microdynamics of civil conflict

During the last decade, a growing number of studies have focused on the microdynamics of conflict, using disaggregate data to analyze conflict variation at a sub-national level (Kalyvas, 2008). This comes partly from a critique to the limitations of the country-year approach that dominated empirical conflict research at the beginning of 2000s (see e.g. Hegre and Sambanis, 2006; Ward et al., 2010). The main idea is to gain a deeper knowledge of how conflict works 'on the ground' and to overcome the drawbacks of data aggregation, which can hide crucial insights about the causes and dynamics of civil conflict.

Accompanied by progress in data collection, several scholars have tried to explain local variations in a single conflict, particularly in relation to different explanatory variables. For example, studies have focused on the effect of commodity price shocks (Dube and Vargas, 2013), inequality (Macours, 2011), interactions among anti-government groups (Metternich et al., 2013), economic development (Barron et al., 2009), ethnicity (Weidmann, 2011), or economic opportunities and greed-related factors (Deininger, 2003) on the spatial or temporal distribution of conflict within a country.

Despite they are not many, some works have applied this same methodology on the Kenyan crisis. Dercon and Gutiérrez-Romero (2012) point out that those households that experienced land disputes and lived in areas where militias were active were more likely to report victimization. Markussen and Mbuvi (2011) suggest that violence is explained by the interaction of ethnic polarization between Kikuyu and the others groups and poverty, youth unemployment, and deteriorated public services. Kasara has used highly disaggregated data in the Rift Valley province to relate conflict intensity to ethnic segregation (Kasara, 2014a) and electoral incentives to redistrict through the use of violence (Kasara, 2014b). The contribution of this article to these works comes from the analysis of violence not only in a spatial disaggregation, but distinguishing as well between organized and non-organized violence, and between incidence and severity.

Disaggregating violence

The disaggregate approach is somewhat linked to a different conceptual understanding of civil conflict. Moving beyond the idea of a civil war with a unique major division and unique interests, conflict is understood as a garbage can of motivations, and outcomes are the consequence of a 'joint production' of violence where several actors with inherently different motivations and opportunities take part (Kalyvas, 2003).

According to this vision, the limitations of data aggregation become evident. Mixing up different types of conflict outcomes in a simple measure means that the data is actually reflecting distinct logics of violence. Violence in riots has a different explanation than violence perpetrated by organized militias, and the logic of participation is completely different for the 'violence entrepreneur' than for the 'plain' members. This articles tries to apply this logic to the Kenyan conflict by proposing a mechanisms that explains the *severity* of *organized* violence. This means that I am using two criteria to distinguish violence. On one hand, violence perpetrated by organized groups is separated from violence that took place in riots. On the other hand, incidence is distinguished from severity, assuming that the decision of organizing violence is not influenced by the same factors that determine its 'success'.

The data on the Kenyan crisis itself points that we should not treat all the violent events as the product of the same process. As figure 1 shows, the initial events were mainly protests or riots where organized groups did not take part.[1] From mid-January onwards, organized groups are the main actors in the conflict. The transition from more spontaneous forms of fighting to organized and planned attacks brings about new actors, new motivations, and new constraints that need to be understood in a different way. Furthermore, figure 2 shows that organized violence is not a mere extension in time, since its spatial distribution is distinct.

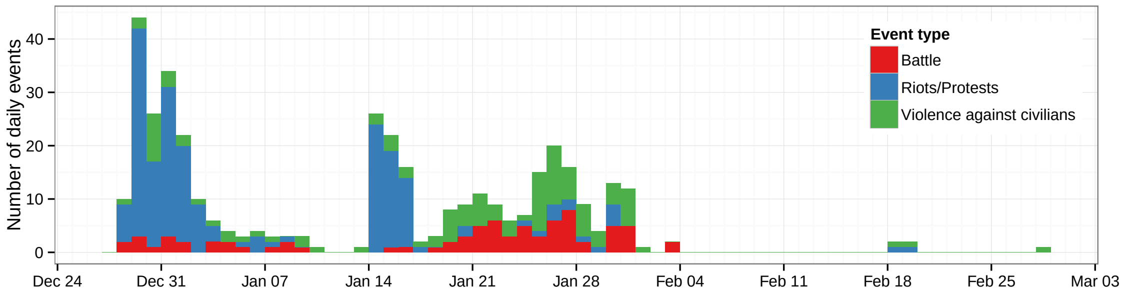


Figure 1. Number of daily events during the 2008 Kenyan crisis.

Source: ACLED (Raleigh et al., 2010).

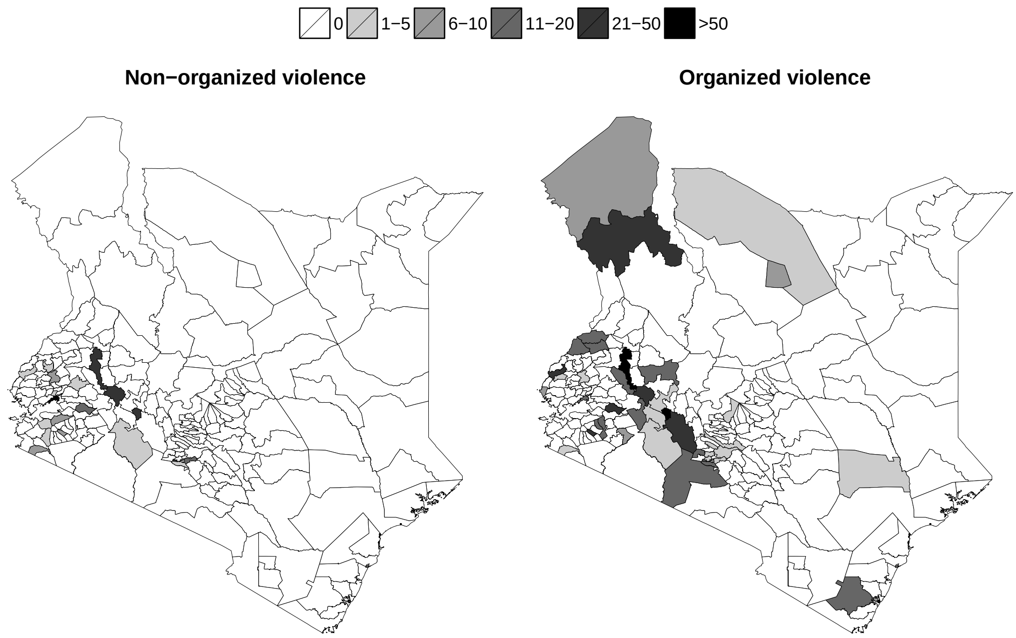


Figure 2. Total fatalities in each constituency in organized and non-organized violence events during the 2008 Kenyan crisis. In ACLED terms, organized violence refers to 'violence against civilians' and 'battles', while non-organized violence corresponds to 'riots/protests'.

Source: ACLED (Raleigh et al., 2010).

Organized violence during the Kenyan crisis

The severity of the organized violence during the Kenyan crisis was the product of a different process than the one explaining its occurrence. While the presence of armed groups depends on the actions of a few leaders that organize the militias, the severity is heavily determined by the success of the recruitment. Here I propose a bottom-up mechanism to explain why the recruitment process was more successful in some areas than in others.

Joining a militia is a risky decision. Participation in ethnic violence may entail negative consequences in the form of eroded social relationships, legal punishment, and physical injuries or death during the clashes. Particularly in the case of Kenya, where the conflict was not part of an ongoing civil war, it is very likely that the individuals assessed these risks when making the decisions of joining an armed group. Indeed, the interviews carried out by Human Rights Watch suggest that the militia members, which were mainly young males, considered both the risks and the material rewards when participating (Human Rights Watch, 2008).

I argue that the decision of joining the groups was heavily determined by the sociopolitical environment in where the would-be members lived. Besides the material rewards for participating, individuals took into account the capacity of leaders to offer legal protection and the likelihood of social sanctions after the end of the conflict. Specifically, a politically homogeneous community where leaders enjoy wide support will facilitate the recruitment process.

First, it is critical that the local leader could offer protection from legal consequences once the conflict have finished. Reports from the early 1990s episodes of ethnic violence in Kenya talk declare the unequal application of the law following political allegiances: 'although the government seemingly has had difficulties in arresting and prosecuting Kalenjin warriors, it has efficiently and quickly prosecuted non-Kalenjins who have acquired weapons to defend themselves after being attacked' (Africa Watch, 1993, 72). Since corruption is widespread in Kenya,[2] potential militia members will perceive lower risks in terms of legal consequences if the local leader has greater political power.

Second, social reactions to the participation itself were part of the pay-off of participating. The public opinion about the use of violence is an important factor behind ethnic clashes, as the participants stay in the community after the conflict and are easily detectable (Horowitz, 2001). As noted in the Human Rights Watch report, 'many of the survivors said many of the attackers were people they knew well' (Human Rights Watch, 2008, 41). A more homogeneous community, not in terms of ethnic divisions but of political allegiances, would reduce the risk of facing social disavowal. Ethnic diversity alone will not affect conflict unless they are articulated along ethnopolitical cleavages.

Thus, the entrepreneurs that decide to organize a militia will be more successful in those places where potential members see them as having enough political resources and where the political homogeneity of the community reduces the risk of social sanctions. These two conditions are found in the party strongholds. However, political hegemony alone will not be a good predictor of violence, since ethnicity was the main driver of clashes and the criteria for selective killings.

Rather, the main condition is the combination between political hegemony *and* ethnic diversity. In other words, violence will be higher where, given some degree of ethnic diversity, ethnopolitical cleavages separate a majority from a minority in terms of political competition.

This argument is tested against the alternative explanation based on the specific rivalry between the Kikuyu and other ethnic groups (e.g. Collier, 2009; Markussen and Mbuvi, 2011). In consequence, the analysis proves that participants' actions were influenced by the sociopolitical context affecting the likelihood of negative consequences, rather that a specific hatred against rival groups.

Data and methods

Data on fatalities

The main data source is the Armed Conflict Location and Event Data Project (ACLED) (Raleigh et al., 2010), a dataset that covers all political violence events reported by media reports, humanitarian organizations, or research publications. The dependent variable is the count of fatalities in each of the 210 Kenyan electoral constituencies from 28th December to 28th February.[3] I only looked at those events where at least one of the parts was an organized armed group. In ACLED coding terms, only 'battle' and 'violence against civilians' events were taken into account. Since the ACLED data is geo-referenced, it is possible to disaggregate the variable to a local level. Besides, the number of total fatalities throughout the conflict measured by ACLED is consistent with other sources, such as the *Waki Comission* report (Waki, 2008).

To distinguish between severity and incidence, I only included in the analysis those constituencies where at least one event of organised violence is reported, even if there were no fatalities. This means that the sample is reduced to 49 cases, of which only four did not witness any death. The appendix presents empirical evidence in favor of the assumption of different generating process for incidence and severity of violence.

Independent variables

Unfortunately, the disaggregation of independent variables to the constituency level was not always possible, which limited considerably the data available. Furthermore, some of the sources used might have some bias, which suggest caution when interpreting the results.

First, the most recent source of publicly available data on ethnicity prior to 2007 in Kenya is the 1989 Census,[4] where it is only possible to calculate ethnic distribution at the district level, using the 1991 administrative division of Kenya in 47 districts. Thus, the ethnic data at each constituency was assumed to be a mirror of the district they were part of. Two variables were created. One measures the degree of ethnic heterogeneity, calculated as $1 - s^2$, where s is the share of the largest group. The other variable indicates the percentage of people belonging to the Kikuyu ethnic group.

Second, the electoral data comes from the Electoral Commission of Kenya, who published the 2007 election results on its website.[5] This data was used to build a measure of political competition, calculating the absolute different in votes between Odinga's ODM and Kibaki's PNU, divided by the total valid votes. Besides, the ODM share of votes in each constituency was also included. As Markussen and Mbuvi (2011) say, this data is not entirely reliable. Since the fraud was widespread, particularly in those areas were the parties had strong support, is very likely that the electoral results are underestimating the actual electoral competition. This fact is directly related to the central argument of this article, however, the analysis was still carried out. The reason is that the mechanism defended here is that violence was higher in the party strongholds, because the strength of the dominant party determined the individual decision of joining a militia and hence the success of the recruitment process. The capacity for altering the election in a given constituency can be seen as well as a measure of political strength, since some kind of resources are needed to do so. Thus, the possible bias in the data is not altering the fundamental inferences about the causal mechanisms.

Finally, economic data was gathered from the Constituency Report of Well-Being in Kenya (Kenya National Bureau of Statistics, 2008), where the data from the Kenyan Integrated Household Budget Survey of 2005-06 (KIHBS) was exploited to estimate poverty measures at the constituency level. The ķihbs is designed to be representative at the district level, and the geographical location of each household is not publicly available, so it is impossible to create disaggregate measures of other economic dimensions. Hence, I included a measure of poverty and the change in the level of poverty between 1999 and 2005-06 in the analysis.[6]

Models

As said before, I create to sets of models to test the suggested mechanism against the Kikuyu-based explanation. On one hand, the Kikuyu variable was hypothesized to be the main factor behind the violence, and was included as a quadratic term (Kikuyu model), to test the effect of the polarization between this ethnic group and the rest of the population, and as an interaction with the ethnic heterogeneity index (Kikuyu model 2). The share of ODM in the elections was also included in these models, to account for the support of Kibaki's opponent. On the other hand, I modeled the dependent variables as a product of the interaction effect between ethnic heterogeneity and political competition (political competition model). Besides, the change in poverty and the population of each constituency (as an offset) are included in all models. The variables of political competition, ethnic heterogeneity, and change in poverty were standardized at $mean = 0$ and $sd = 1$ in the whole sample.

Given the high variance in the dependent variable and its clearly non-normal distribution, even if transformed in logarithmic scale (figure 3), the data was analyzed using negative binomial models, following the same strategy as other works analyzing death counts (e.g. Beger, 2012). To compare the models, I used leave-out-out cross-validation to calculate the root-mean-squared-error (RMSE) and Pearson's R in out-of-sample prediction. This decision follows recent trends in quantitative political research that recommend the use of prediction as a means to test different models (Ward et al., 2010; Hill Jr and Jones, 2014).

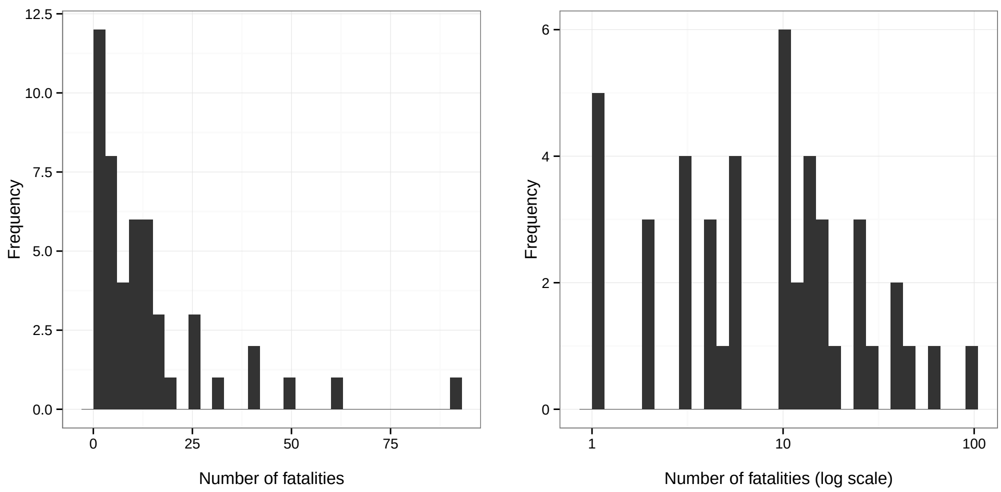
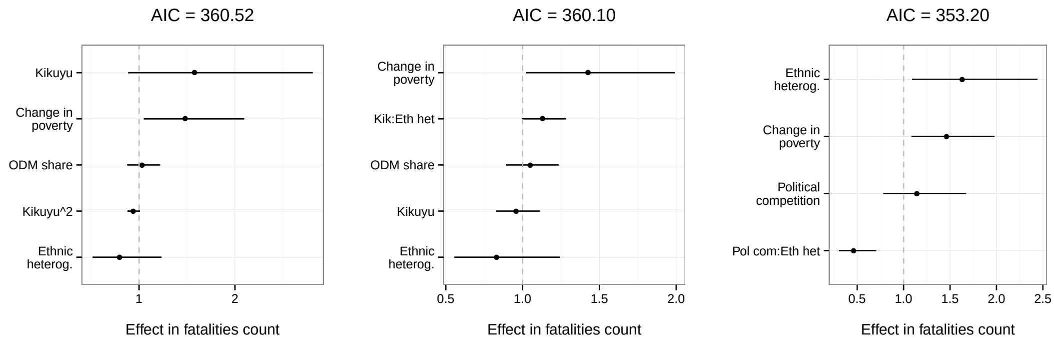


Figure 3. Distribution of the dependent variable, in normal and logarithmic scale.

Results

Figure 4 presents the results of the negative binomial model. Each of the graphs corresponds to one of the three models estimated. The first model (a) includes the Kikuyu variable as a quadratic term, and indicates an effect coherent with the theoretical expectations. The inflexion point is around 40% of Kikuyu population, which means that fatalities are predicted to decrease as the share of Kikuyu group goes away from that point. This is consistent with the findings by Markussen and Mbuvi (2011). The effect, however, is not statistically significant at the 0.05 level. The change in poverty is the only significant variable in the model, and indicates that the increase in poverty between 1999 and 2005/06 is positively correlated with the severity of violence. Each standard deviation above the mean results in an increase of around 50% in the predicted fatalities. The effect of the share of Odinga's ODM is not significant and close to zero.



(a) Kikuyu model (b) Kikuyu model 2 (c) Political competition model

Figure 4. Summary of the negative binomial models. The horizontal lines show the exponential of the coefficient point estimate and its 95% CI in a negative binomial model with the *log* of population as an offset. The sample ($n = 49$) was limited to those constituencies where at least one event of organised violence ('battles' or 'violence against civilians') is reported in the ACLED dataset between 28th December 2007 and 28th February 2008.

The second model (b) introduces the Kikuyu variable interacting with the degree of ethnic heterogeneity. In this case, the only significant variable is the change in poverty, keeping the same effect size as in the previous model (around a 50% increase in deaths count for each standard deviation above the mean).The interaction shows that the negative effect of ethnic heterogeneity is nullified when the Kikuyu group is majority, which is also coherent with the expectation of a quadratic effect of this variable.

The last model (c) tests the main argument of this article. It includes the variable of political competition in a interaction with ethnic heterogeneity, and the change in poverty.[7] Both the ethnic heterogeneity and the change in poverty are significant at the 0.05 level, and present strong effects. Each standard deviation above the mean increases around 65% and 50%, respectively, the expected number of fatalities. However, the effect of ethnic heterogeneity is influenced by the degree of political competition (figure 5), although the low number of cases produces a high standard error. Thus, only in those constituencies where one of the two main parties was dominant, the ethnic heterogeneity was positively correlated with the number of fatalities.

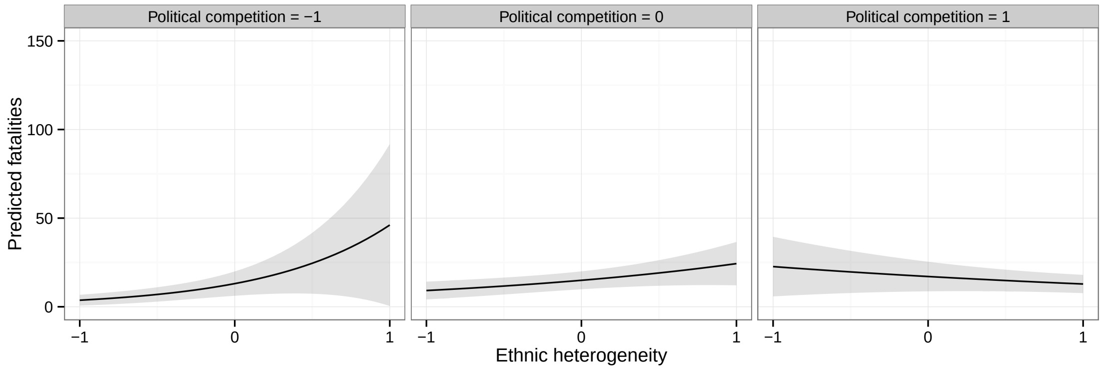


Figure 5. Predicted fatalities at different degrees of ethnic heterogeneity and political competition. The solid line shows the point estimate and the shade area indicates its 95% CI. Both variables are standardised.

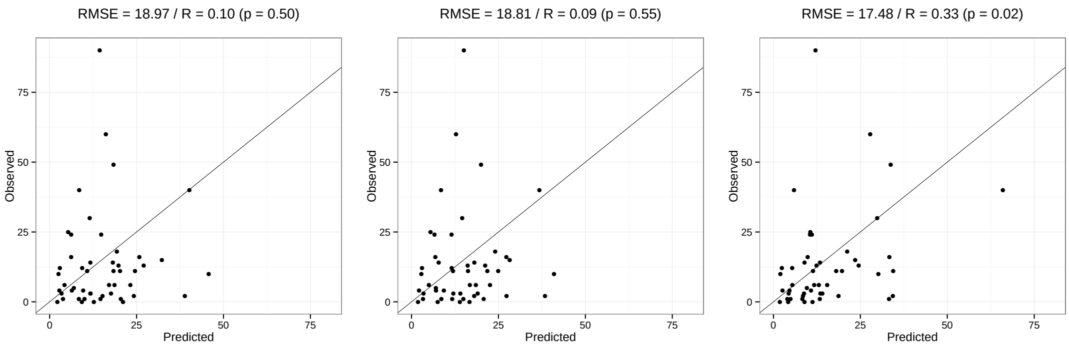
Overall, the change in poverty between 1999 and 2005/06 has a positive (around 50% increase with each stardard deviation) and significant effect in all models. This finding is coherent with past research that highlights the effect of economic dimensions in the intensity of conflict in Kenya (Markussen and Mbuvi, 2011). The fact that the change of poverty has bigger predictive power than the absolute measure of poverty could point that the relationship between economy and conflict is better explained by grievance-based mechanisms.

The figure 6 compares the observed values with the out of sample predictions using leave-one-out cross validation in each of the three models. Both the root mean squared error and the correlation between observed and predicted values indicates that the political competition model presents the best fit. Particularly in some of the regions where violence was more intense, this model is an improvement over the Kikuyu-based explanations. The political competition model also shows the lowest AIC.

Conclusion

Past research on the Kenyan crisis has highlighted the effect of ethnic segregation (Kasara, 2014a), electoral geography (Kasara, 2014b), and the interaction between ethnic tensions and economic problems (Markussen and Mbuvi, 2011) on the severity of violence during the conflict. However, all of them have assumed that every violent event was caused by the same process. In this article I distinguish between organized and non-organized violence, and between severity and incidence of conflict, assuming that these phenomena respond to different process.

I defend that the severity of organized violence was higher in those constituencies where, besides having a high degree of ethnic diversity, one of the



(a) Kikuyu model (b) Kikuyu model 2 (c) Political competition model

Figure 6. Out-of-sample predictive power of the models. The predicted values were estimated using leave-one-out cross-validation. The diagonal line indicates perfect fit between prediction and observations. RMSE indicates the root mean squared error of prediction, and R shows the Pearson correlation coefficient between the observed and predicted values.

main parties was dominant. The idea is that individuals will be more likely to join militias when there is an ethnopolitical majority, because the risk of social sanctions is lower and 'violence entrepreneurs' will have more power to provide legal security.

The empirical analysis presents evidence for this argument. Both in-sample and out-of-sample measures show that the predictive power of the political competition model is stronger than explanations based on the concrete opposition between Kikuyu and the other ethnic groups.

Yet, this work has some limitations. The most important one stems from the unavailability of data at a disaggregated level, that could be biasing the results due to measurement and omitted variable errors. This problem, nonetheless, is common in conflict studies, where fine-grained data on delicate issues is rarely available.

Overall, this article has contributed to the existing literature by stressing the need for a more complex disaggregation of violence, acknowledging the multiple ontology of political violence. Indeed, the disagreement between the findings presented here and past works on the Kenyan crisis might come from the fact that different types of violence are influenced by different factors. Distinguishing between different types of conflict events regarding the different actors and processes is hence critical to a comprehensive understanding of political violence.

Notes

[1] The distinction between 'riots/protests', 'battles', and 'violence against civilians' follows the ACLED coding guidelines. In a nutshell, the difference between 'riots/protests', on one hand, and 'battles' and 'violence against civilians', on the other hand, is the presence or not of organized armed groups. See Raleigh et al. (2014) for further details and definitions.

[2] The country ranked 136th (out of 177) in Transparency International's 2013 Corruption Perception Index, and the problem of corruption is common in the public debate (see e.g. Wrong, 2014)

[3] One case (Kamukunji constituency) is missing due to the non-availability of electoral data. However, there are no recorded events in this constituency.

[4] Using data from the 2009 Census presents important endogeneity problems, given that one of the consequences of the 2008 crisis was a huge amount of internally displaced people (see Markussen and Mbuvi, 2011)

[5] The results were retired after violence erupted, and the Electoral Commission was dissolved since then. I must thank Thomas Markussen for letting me access to the document.

[6] Since both variables are highly correlated (0.7), I decided to use only one of them. The change in poverty performed better in all models (measured in terms of out-of-sample predictive power).

[7] The inclusion of the Kikuyu variable in the model did not present a significant or important effect ($Est. = -0.01, SE = 0.06$) and did not change the effects of the other variables. Besides, it decreased the overall predictive power of the model.

References

Africa Watch (1993) Divide and rule: State-sponsored ethnic violence in Kenya. Report, Human Rights Watch, New York, November.

Barron P, Kaiser K and Pradhan M (2009) Understanding variations in local conflict: Evidence and implications from Indonesia. *World Development* 37(3): 698–713.

Beger A (2012) Using front lines to predict deaths in the Bosnian Civil War. Paper presented at the *International Studies Association 56th Annual Convention*, San Diego, USA, 1-4 April 2012.

Breiman L (2001) Random forests. *Machine Learning* 45(1): 5–32.

Cheeseman N (2008) The 2007 Kenyan election: An introduction. *Journal of Eastern African Studies* 2(2): 166–184.

Collier P (2009) *Wars, guns, and votes: Democracy in dangerous places*. London: Random House.

De Smedt J (2009) 'No Raila, no peace!': Big man politics and election violence at the Kibera grassroots. *African Affairs* 108(433): 581–598.

Deininger K (2003) Causes and consequences of civil strife: Micro-level evidence from Uganda. *Oxford Economic Papers* 55(4): 579–606.

Dercon S and Gutiérrez-Romero R (2012) Triggers and characteristics of the 2007 kenyan electoral violence. *World Development* 40(4): 731–744.

Dube O and Vargas J (2013) Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies* 80(4): 1384–1421.

Hegre H and Sambanis N (2006) Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution* 50(4): 508–535.

Hill Jr DW and Jones ZM (2014) An empirical evaluation of explanations for state repression. *American Political Science Review* 108(3): 661–687.

Horowitz DL (2001) *The deadly ethnic riot*. Berkeley: University of California Press.

Human Rights Watch (2008) Ballots to bullets: Organized political violence and Kenya's crisis of governance. HRW Report, 20 (1), March 2008.

Kalyvas SN (2003) The ontology of 'political violence': Action and identity in civil wars. *Perspective on Politics* 1(3): 475–494.

Kalyvas SN (2008) Promises and pitfalls of an emerging research program: The microdynamics of civil war. In: Kalyvas SN, Shapiro I and Masoud T (eds.), *Order, conflict, and violence*, Cambridge: Cambridge University Press. pp. 397–421.

Kasara K (2014a) Does local ethnic segregation lead to violence? Evidence from Kenya. Unpublished manuscript.

Kasara K (2014b) Electoral geography and conflict: Examining the redistricting through violence in Kenya. Unpublished manuscript.

Kenya National Bureau of Statistics (2008) *Constituency report on well-being in Kenya.* Report, Kenya National Bureau of Statistics, Nairobi.

Macours K (2011) Increasing inequality and civil conflict in nepal. *Oxford Economic Papers* 63(1): 1–26.

Markussen T and Mbuvi K (2011) When does ethnic diversity lead to violence? Evidence from the 2007 elections in Kenya. Department of Economics Discussion Papers 11–19, University of Copenhague, Denmark.

Metternich NW, Dorff C, Gallop M, Weschle S and Ward MD (2013) Anti-government networks in civil conflicts: How network structures affect conflictual behavior. *American Journal of Political Science* 57(4): 892–911.

Mueller SD (2008) The political economy of Kenya's crisis. *Journal of Eastern African Studies* 2(2): 185–210.

Raleigh C, Linke A and Down C (2014) ACLED Codebook 3. Available at: http://bit.ly/1wmkdNt (accessed 12 November 2014).

Raleigh C, Linke A, Hegre H and Karlsen J (2010) Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research* 47(5): 651–660.

Waki PN (2008) *Report of the Commission of Inquiry into Post-Election Violence (CIPEV).* Nairobi: The Government Printer.

Ward MD, Greenhill BD and Bakke KM (2010) The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47(4): 363–375.

Weidmann NB (2011) Violence 'from above' or 'from below'? The role of ethnicity in Bosnia's civil war. *Journal of Politics* 73(4): 1178–1190.

Wrong M (2014) 'Everyone is corrupt in Kenya, even grandmothers'. *Foreign Policy*. May 6. Available at: http://atfp.co/RmKiPu (accessed 12 November 2014).

Appendix

This sections tries to prove that the incidence of organized violence is explained by a different process than the one that explains its severity. To do so, I have run two analysis on the presence of organized armed groups in each constituency. The binary dependent variable takes a positive value if there is at least one report of 'battle' or 'violence against civilians' in a given constituency. This is the same variable that was used to limit the sample in the main analysis of this article.

First, I use random forest (Breiman, 2001) as a non-parametric approach to evaluate the predictive power of each variable, allowing as well for interactions among the covariates (see Hill Jr and Jones, 2014, for a comprehensive explanation and application of this method). Figure 7 shows the mean decrease in accuracy for each of the variables considered, after permuting across 1 000 trees. The bigger the number, the bigger the loss of predictive power in the trees that do not include that variable, and hence the bigger the importance of that variable.

Clearly, the most important variable in terms of predicting whether any constituency will have any event of organized violence is the share of the Kikuyu ethnic group, followed by ethnic heterogeneity and population. The vote share of Odinga's ODM, the degree of political competition, and the change in poverty have considerable less power that the former variables. If incidence and severity were explained by, at least, similar process, we would expect that political competition had much more predictive power. Indeed, this is coherent with the argument defended in the article, since this variable is the one that influences the decisions of the would-be members.

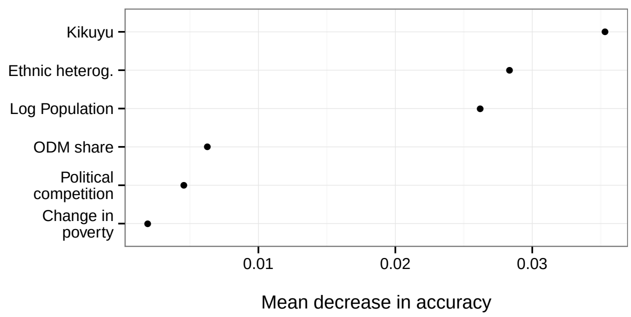
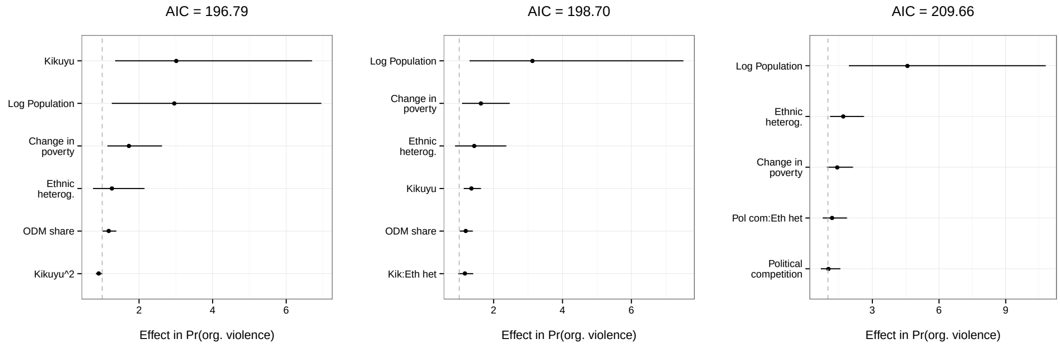


Figure 7. Mean decrease in classification accuracy on whether a given constituency experienced any event of organized violence. The values come from randomly permuting all variables across 1 000 trees, in a random forest analysis using unbiased decision tree algorithm. If a variable is important in explaining the outcome, its exclusion from the trees should bring about a consistent decrease in predictive power.

Second, I use logistic regression to estimate the same models presented in the main text, checking whether the variables affect in the same way the dependent variable. Figure 8 shows the effects of each variable in the three models. Again, it becomes clear that incidence stems from a different causal mechanism than severity. Besides the fact that the relatively goodness-of-fit of the models is completely different, the effects of the variables are not the same.



(a) Kikuyu model (b) Kikuyu model 2 (c) Political competition model

Figure 8. Summary of the logistic regression models on the incidence of organized violence. The horizontal lines show the exponential of the coefficient point estimate and its 95% CI.

Knowing what influences the presence or not of organized armed groups is beyond the scope of this paper. Yet, it becomes crystal clear that incidence cannot be explained by the same causal process than severity.