

Behavioral differences in violence: The case of
intra-group differences of Paramilitaries and
Guerrillas in Colombia

Thomas Bassetti
Raul Caruso
Darwin Cortes

SERIE DOCUMENTOS DE TRABAJO

No. 181

Septiembre de 2015

Behavioral differences in violence: The case of intra-group differences of Paramilitaries and Guerrillas in Colombia

Thomas Bassetti*

Raul Caruso†

Darwin Cortes‡

Abstract

In most studies on civil wars, determinants of conflict have been hitherto explored assuming that actors involved were either unitary or stable. However, if this intra-group homogeneity assumption does not hold, empirical econometric estimates may be biased. We use Fixed Effects Finite Mixture Model (FE-FMM) approach to address this issue that provides a representation of heterogeneity when data originate from different latent classes and the affiliation is unknown. It allows to identify sub-populations within a population as well as the determinants of their behaviors. By combining various data sources for the period 2000-2005, we apply this methodology to the Colombian conflict. Our results highlight a behavioral heterogeneity in guerrilla's armed groups and their distinct economic correlates. By contrast paramilitaries behave as a rather homogenous group.

Keywords: Finite Mixture Models, Colombian Conflict, Behavioral Differences in Conflict

JEL Codes: D74; H72; C46.

* Corresponding author. University of Padua, Department of Economics and Management "Marco Fanno", Via del Santo 33, 35123 Padua, Italy. Email: thomas.bassetti@unipd.it. Phone: +390498274003.

† Catholic University of the Sacred Heart. Via Necchi 5, 20123 Milano. Email: raul.caruso@unicatt.it.

‡ Universidad del Rosario. Facultad de Economía. Calle 12C, No. 4-69, 111711, Bogota, Colombia. Email: darwin.cortes@urosario.edu.co. Earlier versions of this paper have been presented at the 50th Anniversary Conference Conflict Research Society, 2013, University of Essex and at the ICESI workshop on the Colombian conflict, 2014. We warmly thank participants for comments.

1. Introduction

In the latest years, a substantial number of works has analyzed in depth the causes, the determinants and the dynamics of civil wars. However, in most empirical studies, determinants of conflict have been hitherto explored assuming that actors involved were either unitary or stable. In fact, in many internal conflicts a plethora of groups exist or even emerge in the evolution of the struggle. It is not by chance that a growing literature is focusing on this recurring feature of conflicts. Yet, the multifaceted impact of heterogeneity is still to be deepened in both theoretical and empirical studies

In particular, how can we empirically identify different behavioral patterns in multiparty conflicts? Is there an appropriate methodology to detect diversity within groups? Answers to these questions are of great relevance to refine and enrich the knowledge we have about actual conflicts, and in particular civil wars. The existence of unobserved latent groups in a conflict can lead to empirical analyses that under- or overestimate the impact of some important predictors on outcome variables. Therefore, it is particularly important to have a robust identification method that allows us to control for this source of heterogeneity. The aim of this paper is to show how a new model-based clustering technique can be used to tackle this point directly.

The point of departure is that civil wars have to be studied taking into consideration the possibility of a plurality of armed actors. If a plurality of groups exists, it might be argued that they exhibit heterogeneous preferences. Reasonably in the presence of heterogeneous preferences, groups can be also predicted to exhibit different behaviors.

By using longitudinal data, we illustrate how a Fixed Effects Finite Mixture Model (FE-FMM) approach represents an appropriate methodology to explain correlates of different behavioral patterns. This technique allows us to identify sub-populations within a population as well as the determinants of their behaviors. We would be able to figure out the existence of heterogeneity in behavior of different groups and whether or not different groups react differently to specific determinants. Stated differently, whenever we know that multiple armed groups are active in a conflict, but we are unaware on what actions are to be ascribed punctually to each group, by means of a FE-FMM, we can point precisely such behavioral heterogeneity and its determinants.

Needless to say, some theoretical hypotheses have to be available to understand the roots of heterogeneous behavioral patterns. We proceed along the lines of two complementary hypotheses drawn from the existing literature: the competition hypothesis and the inequality hypothesis. The first possible explanation of behavioral differences in violence is the competition between groups. That is, albeit on the same side of a conflict, different groups are likely to compete to increase their status and gain recognition among peers. This inflames conflicts and generates superior violence. In particular, this prediction holds also for subgroups that fight to improve their status within the organization. Heterogeneity in violent events may reflect this. In fact, violence becomes a signal of quality and efficiency.

A second theoretical explanation refers to inequality between groups. On the one hand, inequality has to be interpreted as the inequality in capabilities; and, therefore, outcome of conflict would depend on the capabilities' ratio of parties involved in conflict. On the other hand, inequality may generate a superior grievance in some groups with regard to other parties of the conflict.

We apply this line of reasoning and methodology to Colombian conflict. It is widely known that it has been characterized by the involvement of three violent groups: the right-wing paramilitaries; the Revolutionary Armed Forces of Colombia (*Fuerzas Armadas Revolucionarias de Colombia*, FARC); and the National Liberation Army (*Ejército de Liberación Nacional*, ELN). The Colombian case has received a substantial attention by social scientists [see among others the special issues edited by Brauer & Vargas (2012) and Caruso & Vargas (2014)] but less attention has been paid to behavioral differences between and within armed groups. For this reason, we deepen this aspect in order to verify whether or not it is possible to observe behavioral asymmetries within both paramilitaries and guerrillas. This is particularly relevant in the case of Colombian conflict. In fact, behavioral heterogeneity may contribute to solve the puzzle posed by Rodriguez and Daza (2012) that show how empirical evidence on Colombian conflict is extremely sensitive to the choice of variables, the dataset and the econometric methodology. That is, put differently, scholars are not able to provide an unchallenged evidence on the determinants and the correlates of the conflict so also posing serious concerns in terms of strength of policy implications. In the light of this, unpacking the behavioral heterogeneity of

armed groups serve us faithfully throughout the analysis of the determinants of violence in Colombia.

In the end, we claim several interesting results. First paramilitaries appear to behave as they are a rather homogeneous body. Moreover, when considering correlates of paramilitaries' activity, they are rather sensitive to rents emerging from coca production as well as from oil, metal and coal mining. We have found a link between paramilitaries' activity and coal areas.¹ By contrast, our analysis confirms that *guerrilleros* appear to be heterogeneous. They have to be divided into two different types. In fact, these types differ in terms of both preferences and locations. A first type appears to reflect a clear-cut predatory behavior whereas the second shows a more defensive or conservative behavior. In fact, the first type attacks increase in the presence of both legal and illegal rents whereas the attacks of second type decrease with coca crops (defensive behavior) and are concentrated in less urbanized municipalities (with weaker institutions) in the Andean region (the most developed in the country). Some of these results are in line with the existing literature but at the same time some of them provide a novel evidence on the dynamics of violence in Colombia.

In general, these results confirm that our approach has interesting applications in the realm of positive analysis of conflict. Our study has also some relevant normative aspects. If behavioral differences shape the actual conflict, they also would shape the conflict resolution process. In particular, in the presence of a significant evidence of systematic differences between violent groups, strategies of conflict resolution have to take into account preferences and peculiarities of different groups.

The rest of paper is organized as follows. Sections 2 and 3 present some related literature. Section 4 presents the FE-FMM approach, while Section 5 applies this estimation strategy to the Colombian case. Section 6 concludes.

2. On Multiparty Conflicts

This paper extensively draws insights from works analyzing multi-party conflicts. Traditionally, the study of conflict had focused on dyadic interaction in both interstate and intrastate conflicts, but recently, some

¹ Some judicial inquiries emerged on the so called "blood coal". According to witnesses, mining companies were used to support financially paramilitaries to run operations against guerrillas (see Moor and van de Sandt, 2014).

works have begun analyzing multiparty conflicts and in particular with regard to the severity and duration of a civil conflict and eventually to the impact on stability of peace agreements.

We are interested in the theoretical propositions that explore under what conditions the existence of several armed groups can increase the severity of a conflict. Butler et al. (2014) model a multiparty contest showing that the existence of multiple parties increases the intensity of conflict. These predictions are confirmed by an empirical analysis on internal armed conflicts from 1989 to 2012. Other works focus on asymmetries in wealth, endowments and capabilities. There may be several theoretical explanations on this relationship. For the sake of simplicity, consider an armed group split into two subgroups. They are asymmetric in endowments and economic opportunities. On the one hand, drawing from Gurr (1968), the poorer subgroup may have superior resentment and anger, exerting a higher level of violence. It may have superior incentives to appropriate the largest possible fraction of some appropriable income stream or assets. That is, poorer parties in conflicts have the incentive to fight more aggressively in order to improve the current income distribution. This was labeled as the paradox of power by Hirshleifer (1991). Esteban and Ray (2011) present a model of conflict which addresses the role of within-group inequality, showing that within-group inequality increases conflict. In their model, money and time are treated as substitutes. In fact, members of the group can contribute with either time or money to conflict. They show that a superior degree of within-group inequality is likely to make groups more violent because the opportunity cost of time for poorer member decreases. Activism of members is essentially mercenary so richer members can exchange money with higher levels of activism from poorer members of the group. Cubel and Sanchez-Pages (2012) present a model that strengthens this proposition. They show that the technology of fighting plays a crucial role in the presence of unequal endowments within-groups. A different theoretical explanation on why the existence of multiple groups increases the severity of conflict is in Munster and Klaas (2011). The authors study analytically the interdependence of internal conflict and external conflict by means of a model in which players are partitioned in groups and have to choose whether to fight against other groups or within their own group. Under some conditions, groups are to choose external conflict rather than internal conflict. That is, external conflict is motivated by willingness to predate outside rather than fighting internally for appropriation. Yet, when

power is insecure within groups, the model also predicts that an internal fighting will occur. In recent years, several empirical studies also have analyzed the impact of multiple actors on conflicts. Cunningham (2013) explores to which extent the existence of faction in self-determination movements is associated with a greater probability of a civil war. The empirical analysis covers one thousand and one hundred eighty-eight factions within one hundred and forty six movements in seventy-seven countries. Results highlight that more fragmented movements, namely those with a larger number of internal factions are more likely to be involved in a civil war.

Other studies test the hypothesis of competition between groups as source of higher severity. According to Gupta and Mundra (2005), the number of Palestinian suicide attacks in the period 1991-2003 is the result of both cooperation and competition among Palestinian groups. More specifically, previous Palestine Liberation Organization's incidents cause current attacks by Hamas and Islamic Jihad, while previous attacks by Hamas induce also attacks perpetrated by Islamic Jihad. At the same time, Palestinian suicide attacks are also a reaction to Israeli attacks. Similar explanations have been provided by Bloom (2004) and Della Porta (1995). Bloom studies suicide bombing by Palestinian militants, while Della Porta considers the competition among different terrorist groups in Italy in 1970s. Clauzet et al. (2010) also analyze in depth the dynamics of inter-group competition between Palestinian groups. The authors confirm that behavior of Palestinian groups is by no means only dependent upon Israeli strategy. In fact, it appears that Palestinian groups' actions are motivated by both Israeli strategy and dynamics between other Palestinian groups. Findley and Young (2012) also explore the same hypothesis finding support for suicide terrorism. Caruso and Schneider (2013) found that a competition between Islamist groups may explain the upward brutality of terrorist incidents.

In addition, there is evidence that outcome of conflicts is severely affected by the existence of multiple groups. Peace agreements are, in fact, fragile and such fragility strongly depends on the presence of different groups (Nilsson, 2008, 2010; Bove and Smith, 2011). Furthermore, duration of civil conflict has been found positively associated with the number of violent groups. (see, among others, Cunningham, 2006; Cunningham, 2010; Akcinaroglu, 2012).

3. A Brief History of the Colombian Conflict

The origin of Colombian conflict can be traced back to the period known as *La Violencia*. *La Violencia* begun in 1948 with the assassination of Jorge Eliécer Gaitán. Gaitán, a populist leader of the Liberal Party and the party candidate to the presidency of the republic. His death generated an escalation of violence between the two principal political parties: the Liberal Party and the Conservative Party. This period ended in 1957, when general Gustavo Rojas Pinilla allowed Liberals and Conservatives to reach the National Front Agreement.² However, Rojas Pinilla's regime that united the two parties had a strong opposition (Offstein, 2003). During the National Front the Liberal and Conservative parties monopolized the Colombian political system for 16 years.³

In 1963, inspired by Fidel Castro's revolution, peasants, Catholic radicals, students and left-wing intellectuals constituted ELN. One year later, communist militants and peasants decided to form FARC. Both groups were originated by the conflict in "*La Violencia*" and the fact that other political positions different from those in the Liberal and Conservative parties were excluded from the political system. Both groups represent rural peasants and colons and oppose rich classes and US interference in internal affairs. In some areas of the country and periods, these two groups have cooperated, but they have had some armed clashes in other areas and years.

FARC and ELN have three main differences (Medina, 2010). First, FARC always had a political party to express its political positions. Under the strategy of using "all forms of fight combined", FARC could express its political positions through the Colombian Communist Party (CCP). Differently, ELN did not have this type of links with the political system and political positions were expressed through social movements. Second, the relationship between FARC and illegal drug trade goes back to early 1980s. This might be generated profound differences between rich and poor groups within FARC. Instead, ELN committed to not be involved in drug trade. This commitment lasted until early 1990s when ELN entered into these predatory activities. Third, FARC initiated peace processes earlier than ELN. Effectively, in middle 1980s FARC did a process with Betancur government. This process ended up with the creation of a new FARC political party ("the Patriotic Union", UP) that signaled the end of FARC

² General Gustavo Rojas Pinilla was responsible for the coup d'etat that ousted Laureano Gómez in 1953.

³ The Agreement consisted in that these two parties alternated in power each four years.

relationship with the CCP. Afterwards, paramilitaries in coalition with the Colombian army killed most of UP militants, so that by the end of 1980s FARC did not have a political party anymore. That is, at the same time that FARC had access to huge revenues from drug trade and increased its military strength suddenly had no room to express its political positions. By the end of 1990s both FARC and ELN being illegal armies that predate resources, FARC had a much more strong army than ELN, but with less chances to express its political positions. Somehow, the 1998-2002 peace process gave some opportunities for FARC to express its political positions (through the Clandestine Colombian Communist Party and the Bolivarian Movement), but large internal differences were already in place.

On the other hand, the United Self-Defense Forces of Colombia (Autodefensas Unidas de Colombia, AUC) is the coalition formed by most of local paramilitary groups. The coalition was created in 1997 to consolidate self-defense groups that were created during the 1980s to defend the interests of landlords and drug barons from guerrilla. However, in line with the legislation, paramilitaries already existed as self-defense groups since the 1960s. That is, they aim to protect dominant class's interests. In 2003, there was a partial and polemical demobilization process of the AUC. The maximum leaders of the AUC were extradited to the US under narcotrafficking charges, leaving the demobilization process unfinished. In 2005, new paramilitary groups were formed by new combatants and factions of the former AUC groups.

The analysis of the Colombian conflict has received a substantial attention by economists. Vargas (2012) shows that conflict persistence increases with the presence of illegal rents from coca crops and legal rents from mining. Besides, the author shows that the stronger institutions and the more active military, the shorter the conflict. The author makes no claim on interpreting these associations as causes but it is useful to summarize the posterior work that to some extent explores the highlighted relationships. The analysis of causes of Colombian conflict involves several pieces of the literature. Firstly, regarding the analysis of illegal rents, Angrist and Kugler (2008) show that rents from coca exacerbate homicides in regions where guerrilla was already in place. The analysis of illegal rents and the Colombian conflict is continued by Idrobo et al. (2014), who show that rents from illegal gold mining increase homicides and massacres from both illegal armed groups, but do not exacerbate displacement. The interpretation of authors is that the presence of illegal gold mining

opportunities makes illegal armies to fight each other for the control of these regions. But, since illegal gold mining is intensive on the use of labor, it does not cause forced displacement. Secondly, in the analysis of legal rents, Dube and Vargas (2013) show that (positive) income effects might have opposite effects depending on technology production. The authors show that legal rents might either mitigate conflict when they are associated to the production of labor intensive goods (e.g. coffee) or exacerbate conflict when they are associated to the production of capital intensive goods (e.g. oil). The authors show that a fall of coffee prices exacerbates both guerrilla and paramilitary attacks (the opportunity cost effect). In contrast, a rise in oil prices only increases paramilitary attacks (the predatory effect). The role of production technology is further analyzed by Palacios (2012), who finds larger displacement in the presence of economic activities that uses labor less intensively. The role of military presence is further analyzed in Cortes et al., (2012) that show a non-linear effect of police on conflict. Initial police deployments increase guerrilla attacks and posterior police reinforces decrease guerrilla attacks. Police policies have no effect on paramilitary attacks. Finally, Cortes et al. (2014) show that a larger provision of roads helps to mitigate the onset of conflict of both guerrilla and paramilitary but only affects the intensity of paramilitary attacks.

To sum up, provided that the presence of coca crops and other illegal or legal activities (and the technology they use) varies across Colombian geography as well as the strength of institutions and the presence of the legal military forces, this might call for different behavior/strategies of illegal armies in different regions.⁴

4. Data, methodology and estimations

⁴ This paper relates mostly to the literature on determinants and causes of Colombian conflict. However, the literature on both economic and political consequences of conflict is also large. On economic grounds, it has been shown that actions of both guerrilla and paramilitary decreases social development (Diaz and Sanchez, 2004), and education (Barrera and Ibáñez, 2004; Dueñas and Sánchez, 2007; Rodríguez and Sánchez, 2009, 2010); and increases poverty (Lemus, 2014). On political grounds, Acemoglu et al. (2010) explain the connivance of politicians with paramilitary troops trough an electoral mechanism. Moreover, Gallego (2011) shows evidence that support the idea that the guerrilla tries to sabotage elections and paramilitary create alliances with certain candidates. None of these contributions look at differences of behavior within illegal groups.

4.1 The data

Data come from several sources. The descriptive statistics are summarized in Table 1. The conflict variables come from CERAC, a Colombian NGO specialized in the analysis of the Colombian conflict. Their database is based on the number of events committed by the actors involved in the conflict. This data is used in most of papers mentioned in the literature on Colombian conflict and is explained with detail in Restrepo et al. (2004).⁵

As dependent variables, we use the number of guerrilla and paramilitary unilateral attacks per ten square km perpetrated yearly for the period 1999-2005 at the municipality level. The average incidence of attacks of both groups is quite small (less than one attack per year per municipality). The most attacked municipalities suffered up to 30 guerrilla attacks a year and up to 9 paramilitary attacks a year, showing that attacks are highly concentrated in a relatively small number of municipalities. Since municipalities differ significantly in terms of extension of the territory, we calculate the number of attacks per ten square kilometers in order to have a comparable outcome for all municipalities. On average, the guerrilla attacks more frequently than the paramilitaries.⁶ Since both groups exhibit a highly skewed distribution of the outcome, we transform our dependent variables by taking the natural logarithms.

To avoid endogeneity issue following Gallego (2011), we instrument paramilitary and guerrilla attacks using geographic variables. In particular, we use the Euclidean distance in kilometers of municipalities to San Vicente del Caguán (also known as El Caguán) as exogenous source of variation in guerrillas' attacks. Between 1999 and 2001, President Pastrana demilitarized the area around El Caguán in order to facilitate the peace dialogue. During this period, FARC's used this territory for training and recruiting, continuing to perpetrate a series of crimes such as drug trafficking, extortion, kidnappings, etc. At the beginning of 2001, Pastrana's government declared the failure of negotiations and attempted to recover these areas. Obviously, this attempt led to a dramatic increase in the conflict especially around the demilitarized zone. Therefore, we use the

⁵ Using two different conflict data, Rodriguez et al. (2012) show that no determinant is robust to different specifications. We use CERAC data that is used, for instance, in Dube et al. (2013).

⁶ The reader should be aware that the Colombian conflict is the consequence of three different forces: paramilitaries, guerrillas and legal army. Therefore, a complete analysis should also consider the number of attacks launched by the government. However, a three-part conflict would extremely complicate the analysis. For this reason, we proxy government intervention with State capacity variables.

interaction between the distance to El Caguán with the temporal distance to the end of the negotiates as IV for FARC's attacks. Since paramilitaries were not part of the peace dialogue, there is no specific reason why this variable should directly affect paramilitaries' attacks.

The instrument for paramilitaries' attacks is based on the so called *Demobilization Program* implemented by President Uribe. Between 2003 and February 2006, about 17,000 of the AUC's 20,000 fighters surrendered their weapons. As argued by Mingorance et al. (2004) the paramilitaries' reintegration program was based on the cultivation of oil palm trees. This led to an increasing interest of paramilitaries for areas suitable for the cultivation of palm trees.⁷ Therefore, we instrument paramilitaries' attacks with an interaction term between a dummy variable indicating whether palm trees can be cultivated in a municipality and a dummy variable indicating the demobilization period. No evident reason justifies a direct association between the possible cultivation of palm trees and guerrilla warfare. On the contrary, the project of paramilitary expansion is correlated with the cultivation of oil palm trees (Gallego, 2011).

In the analysis, we control for both time-varying and time-invariant regressors. However, when we adopt a fixed effects specification the effects of time-invariant regressors and unobservable characteristics are completely absorbed by the individual component. This also excludes that our results are capturing geographic determinants instead of behavioral ones. Time-varying regressors include several dimensions. First, we control for the population size. This variable comes from the Colombian National Statistics Office (DANE from its acronym in Spanish). The economic variables have two sources. First we exploit a composite index, *PC economy*, made of fiscal data reported by the municipality governments to the National Department of Planning. This index is the first principal component computed using total income, tax revenue, national transfers received by the municipality, expenditure, and investment.⁸ All variables are measured in per capita terms. The Gini index is calculated by the DANE. Table 1 shows that municipalities heavily differ with respect to poverty and income inequality.

⁷ Oil palm trees need an altitude below 500 meters above the sea level and an average temperature between 22 and 32 degrees Celsius (see Mingorance et al., 2004).

⁸ In fact, including collinear variable would lead to imprecise estimates; at the same time, excluding relevant variables from the model can result in bias. We use a principal component analysis to summarize correlated indicators mediating between these two problems. The largest eigenvalue is 3.897, while the second largest eigenvalue is 0.592. Therefore, using 1 as a rule of thumb to choose the number of components, it is clear that the first component already summarizes the information we need.

We also control for the availability of natural resources. The variables on coca production come from the Integrated Illicit Crops Monitoring System (SIMCI) that is managed by the United Nations Office on Drugs and Crime (UNODC) in Colombia. This system uses satellite images information that is corroborated on ground investigation. The SIMCI monitors the territory continuously such that the data changes both across municipalities and time. We use two variables: *Coca* and *Coca × North*. The former indicates the number of hectares cultivated with coca leaf, while the latter is obtained by interacting *Coca* with a dummy variable indicating whether a municipality is located at the north of Bogota (a region recognized to be under the influence of paramilitaries). The variables on coffee, coal, oil and metals (gold and nickel) come from the DANE. They estimate the proportion (in percentage) of the production of these resources on the total GDP of the department. We use the ratio between this proportion and the department area (in square Km). Coal and oil production is concentrated in a small quantity of departments.

When we use a random effects approach, we also include a number of time-invariant regressors. To capture state capacity and credit institutions (banks) in the municipality, we include different variables of institutions.⁹ They consist of the number of institutions per one thousand inhabitants in 1995 at the municipality level. The source is the Fundación Social, a Jesuit NGO based in Colombia. The security institutions are police stations. The law enforcement institutions are tribunals, notaries and jails. The financial institutions are different types of commercial banks. The social institutions are hospitals, health centers and schools. The bureaucracy institutions are offices of Public Instruments Registry and tax collection offices. Regarding institutions, disparities across municipalities are correlated to social institutions, where some municipalities have no social institutions. The geographical variables come from the Colombian National Institute of Geography (IGAC) and the Colombian National Institute of Weather (IDEAM). The distance to markets is the average Euclidian distance to the main four markets, namely, Bogota, Medellín, Cali and Barranquilla. It is measured in thousands of kilometers. Rainfall corresponds to the yearly average of precipitations in millimeters. We also control for the percentage of population with unmet basic needs from the 1993 population census.

Table 1: Descriptive statistics (5303 observations, 960 municipalities)

⁹ It is worth noting that, in most municipalities banks are owned by the State itself.

Variable	Mean	Std. Dev.	Min	Max
Dependent variables and instruments				
Guerrillas attacks	0.679	1.836	0	30
Paramilitaries attacks	0.141	0.581	0	9
Area	0.727	1.790	0.018	18.381
IV (Gue)	315.182	949.530	-2374.078	3561.117
IV (Para)	0.117	0.321	0	1
Time-varying regressors				
Coca	74.461	634.492	0	16524.000
Coca × North	17.926	223.687	0	10566.000
PC Economy	0.007	2.127	-0.116	118.641
Gini	0.692	0.111	0.301	1.065
Coffee	0.027	0.076	0	0.843
Coal	0.007	0.037	0	0.345
Oil	0.014	0.041	0	0.350
Metal	0.004	0.013	0	0.092
Population	0.415	2.548	0.011	70.299
Time- invariant regressors				
Security	0.156	0.126	0	1.138
Law enf.	0.243	0.196	0	1.706
Financial	0.199	0.158	0	1.363
Social	0.542	0.411	0	9.990
Bureaucracy	0.049	0.105	0	1.986
Dist. Mkts	0.329	0.138	0.164	1.094
Poors 93	51.555	17.963	10.7877	100
Rainfall	1876.151	1024.215	160	9200

Note: Guerrilla and Paramilitary attacks are measured as the number of unilateral attacks per ten square kilometers. IV (Gue) is the Euclidean distance between the centroids of each municipality and San Vicente del Caguan, the municipality at the heart of the demilitarized zone for the peace process held between 1998 and 2002. IV (Para) is a variable that, since 2003, takes value one for those municipalities under 500 meters above the sea level. Institutions are measured at 1995. Coca prod is the number of hectares cultivated with coca leaf. Coca north is the same as Coca prod, for municipalities at the north of Bogota. PC Economy is the first principal component of a vector of variables related to the municipality economy, namely, total income of municipality government, tax revenue, national transfers, expenditure and investment of municipality government. Gini is the income Gini index. Dist. markets is the average Euclidian distance to the main four markets of the country, namely, Bogota, Medellin, Cali and Barranquilla. Area is measured in thousand of squared kilometers. Coffee, coal, oil and metal areas are the proportion of the production of these resources on the total GDP of the department. Security institutions are police stations per one thousand inhabitants. Law enforcement institutions are tribunals, notaries and jails per one thousand inhabitants. Financial institutions are different types of commercial banks per one thousand inhabitants. Social institutions are hospitals, health centers and schools per one thousand inhabitants. Bureaucracy

4.2 A general analysis of the Colombian conflict

By exploiting the panel structure of our data set, we first estimate two standard linear models: a random effects (RE) model and a fixed effects (FE) model. In particular, we estimate the following specification:

$$y_{it}^j = d_i + x_{it}'\beta + \varepsilon_{it}, \quad (1)$$

where y_{it}^j is the number of attacks undertaken by group j in area i at time t divided by the area of municipality i , d_i is the area specific effect (random or fixed), x_{it} is the vector of explanatory variables, β is the vector of parameters and $\varepsilon_{it} \sim N(0, \sigma^2)$ is the error term. For the sake of notation, we omit the index j when there are no ambiguities. The estimated parameters are reported in Table 2.

Table 2: RE and FE estimates for guerrillas and paramilitaries

	Guerrillas		Paramilitaries	
	RE	FE	RE	FE
	(1)	(2)	(3)	(4)
Rival group	2.0770*** (55.043)	2.1661*** (50.670)	16.9732*** (68.966)	18.1086*** (103.134)
FS residual	-1.9806*** (353.324)	-1.9846*** (341.560)	-16.0226*** (71.328)	-17.1102*** (116.400)
Coca	-0.0001*** (9.426)	-0.0002*** (8.651)	0.0019*** (54.774)	0.0020*** (55.339)
Coca \times North	0.0002 (0.705)	0.0004** (2.516)	-0.0005*** (3.457)	-0.0005*** (3.102)
Economy	0.0099*** (3.166)	0.0114*** (3.291)	0.0034 (0.532)	0.0025 (0.369)
Gini	0.7518*** (7.676)	0.4701 (0.392)	20.6704*** (51.489)	19.3386*** (10.587)
Coffee	0.6488*** (4.544)	0.7410*** (3.917)	-16.9321*** (29.453)	-18.0512*** (27.085)
Coal	2.9568*** (7.715)	12.5132*** (13.059)	64.5168*** (47.624)	81.3435*** (45.974)
Oil	10.4671*** (14.600)	7.4337*** (7.264)	76.7500*** (67.131)	75.2590*** (49.161)
Metal	14.1883*** (3.582)	38.2493*** (7.627)	298.4885*** (57.417)	374.8775*** (47.073)
Population	0.0924*** (4.500)	0.0675 (0.575)	0.2308*** (3.538)	1.4018*** (7.010)
Security	-2.3789*** (7.510)		-22.5300*** (36.685)	

Law enf.	-1.2496*** (16.944)		0.2583 (1.357)	
Financial	-1.3270*** (3.308)		-17.4030*** (19.939)	
Social	0.6284*** (20.667)		-2.8927*** (53.701)	
Bureaucracy	3.7081*** (23.426)		7.7310*** (21.769)	
Distance to mkts	3.5244*** (23.895)		23.7923*** (45.375)	
Poors 93	0.0260*** (22.223)		0.2048*** (69.482)	
Rainfall	0.0004*** (8.041)		0.0029*** (61.100)	
Mills' ratio	0.9279** (2.444)	1.8205*** (4.287)	3.1465*** (11.361)	3.2918*** (9.142)
North	-0.0165 (0.942)		-0.0331 (0.748)	
Constant	-10.0134*** (16.060)	-8.3617*** (8.042)	-139.4890*** (67.205)	-130.8329*** (58.000)
Observations	5,303	5,303	5,303	5,303
Number of munic.		960	960	960
R-squared		0.981		0.891
Robust z-statistics in parentheses. Significant at: *10%, **5%, ***1%.				

Columns 1 and 2 of Table 2 show that the number of attacks undertaken by paramilitaries has a positive impact on the number of attacks undertaken by guerrillas. This means that guerrillas respond to paramilitaries' attacks increasing the number of their offenses. There is a negative correlation between guerrillas' attacks and the production of coca in municipalities located at the South of Bogota.¹⁰ Besides, the higher the economic development, the higher the number of guerrilla attacks. Column 3 also shows the existence of a positive association between the number of attacks undertaken by guerrillas and the availability of important resources such as coffee, coal, oil and metals. The positive impact of oil production on guerrillas' attacks captures an important feature of the conflict. At the beginning of the twenty-first century, rebels surrounded oil fields and the petroleum industry was at a standstill. The relationship between guerrillas and oil companies was complex. There is a huge record of killings and kidnappings of oil workers. Moreover, rebels used to target oil sector first destructing oil pipelines and then asking for protection fees from firms.

¹⁰ The coefficient of *Coca* measures the impact of coca production in southern areas, whereas the effects of coca production in northern areas is the sum of the two coefficients.

Insurgents nowadays continue to attack areas characterized by oil production.

Notice that, even if in the RE specification we control for several time-invariant regressors, these estimates are inconsistent. This conclusion is confirmed by a Hausman specification test. RE and FE estimates for the number of attacks undertaken by paramilitaries are presented in Columns 3 and 4 of Table 2, respectively. Paramilitaries' attacks are positively associated with guerrillas' attacks. Columns 3 and 4 suggest that estimates in Columns 1 and 2 are affected by a simultaneity bias due to the fact that paramilitaries respond to guerrillas' attacks and this inflates the conflict. There is a positive correlation between coca exploitation and paramilitaries' actions. Even in the northern areas, although the magnitude is smaller. Paramilitaries are active in populated municipalities in which the distribution of wealth is more unequal. These regions are mainly urban and suburban areas. In contrast, paramilitaries seem not to be interested in rural areas characterized by coffee production. Our estimates also show that, between 1999 and 2005, paramilitaries operated especially in municipalities characterized by the presence of mining sites. In line with Dube and Vargas (2013), we find that paramilitaries pay attention to oil and gas areas. According to Sanchez and Palau (2006), this predatory activity increased after the fiscal decentralization put in place in 1991. Moor and van de Sandt (2014) argue that, during this period, some coal mining companies in Colombia have supported paramilitary groups to protect their interests. Victims and perpetrators have reported that companies provided material support and strategic information to paramilitaries in order to free the mining region of rebels. Given the correlation between unobserved characteristics and explanatory variables, a FE approach is more appropriate than a RE approach.

4.3 Detecting heterogeneity by means of Finite Mixture Models

Results reported in Table 2 provide a clear and reasonable picture of the Colombian conflict. However, so far, we have treated paramilitaries and guerrillas as two homogeneous groups, ignoring the existence of latent classes within each group. We address this issue by exploring intra-group heterogeneity. In what follows, we present the Finite Mixture Models

(FMM) and eventually we compare the results with those of the previous section.

As noted above the FMM provide a representation of heterogeneity when data originate from different latent classes and the affiliation is unknown. In general, a FMM concerns modeling a certain distribution by a weighted sum of other distributions and is based on the following assumptions:

- i. Each observation is assumed to be distributed according to a mixture of K different components (groups).
- ii. Each component comes from the same family of distributions, but these distributions have different parameters.

Recently, FMMs have been extended to panel data analyses (see, e.g., Skrondal and Rabe-Hesketh, 2004; Bago d'Uva, 2005; Deb and Trivedi, 2013). In particular, considering a short panel set-up, Deb and Trivedi (2013) developed finite mixture models with fixed effects (FE-FMM) for two families of distributions: mixtures of Normals and mixtures of Poisson. In this way, the correlation between the unobserved individual heterogeneity and the regressors does not affect the identification of latent classes. Since a mixture of Normals with fixed effects can be estimated through a standard FMM for cross-section data once the within transformation to the data has been applied, we use this technique to determine the optimal number of components. Under normality assumption, we can write the conditional likelihood for the k -component of the mixture model as follows:

$$f_k(\beta_k, \sigma_k^2 | d) = \prod_{i=1}^N \frac{(2\pi\sigma_k^2)^{-T/2}}{(2\pi\sigma_k^2/T)^{-1/2}} \exp\left\{\sum_{t=1}^T -\frac{1}{2\sigma_k^2} [(y_{it} - \bar{y}_i) - (x'_{it} - \bar{x}_i')\beta]^2\right\}, \quad (2)$$

where \bar{y}_i and \bar{x}_i are the individual means, N is the cross-section dimension, T is the time dimension, $d = T\bar{y}_i$ is a sufficient statistic for d_i . By denoting with π_k the posterior probability of component k , the mixture distribution can be written as

$$\sum_{k=1}^K \pi_k f_k(\beta_k, \sigma_k^2 | \bar{y}_i, x_{it}) \quad (3)$$

with $\pi_k \in (0,1)$ for any $k=1,\dots,K$, and $\sum_{k=1}^K \pi_k = 1$. The vector of parameters is obtained by maximizing the full-data likelihood function.¹¹ Because ex-ante we do not know whether a data point belongs to a subgroup or another, we use the famous expectation-maximization (EM) algorithm designed by Dempster et al. (1977).

On the basis of posterior probabilities, we assign each observation to the subgroup with maximum posterior probability. This is the traditional way in which FMMs are used as intrinsic classification models. We measure the quality of our classification by using an entropy index. This index ranges from 0 to 1 and measures the classification accuracy of a mixture model. Higher values indicate a better classification (see Celeux and Soromenho, 1996).

An important estimation issue concerns the choice of K . In fact, the number of components must be specified before we start the EM algorithm, and we cannot rely on any theoretical insight. Therefore, even if in general few components are sufficient to fit density well, we use the usual information criteria to determine the optimal number of components. More specifically, starting from $K=2$ and increasing the number of components, we pick the model with the largest entropy index and the smallest information criteria (AIC, BIC and sample-size adjusted BIC). If the entropy index measures the distance among groups, information criteria measure the performance of model fit. However, as a rule of thumb, we reject any mixture model with an entropy index lower than 0.6. When the entropy index is less than 0.6, the fraction of misclassified observations is more than 20% (see Lubke and Muthén, 2007).

To avoid any endogeneity bias we use a control function (CF) method to address endogeneity issues. By using an excluded instrument, we estimate residuals from a first stage regression and we insert them into the second stage. The intuition behind the CF method is the following: if a covariate is affected by the outcome variable, regressing this covariate on a set of exogenous regressors would lead to an error term correlated with the initial outcome; therefore, including this error term in the second-stage equation will lead to less biased estimates. For linear models, the control function method mimics a two-stage least squares approach, whereas for more complex models the former offers additional advantages (see Cameron and Trivedi, 2013). Both methods rely on the identification of a proper

¹¹ See Deb and Trivedi (2013) for further details on the maximization algorithm. The estimates have been produced using Stata (version 13).

instrument. Finally, the inverse Mills ratio can be used to control for censoring problems generated by the fact that groups often target specific areas.

Table 3 below reports the statistics we used to determine the optimal number of mixture components. Since the FE-FMM for paramilitaries does not converge to a suitable mixture, we explore guerrillas' behavioral differences, assuming that paramilitaries act as an homogenous group. As we already said, the decision rule is based on information criteria and a measure of entropy. AIC and BIC describe the goodness of the fit and are more reliable than the maximal likelihood, which would lead to over-fitted solutions. However, these statistics do not take into account the distance among components. For this reason, we also consider an entropy index, which measures the capacity of the model to identify the latent classes, i.e., the quality of the mixture.

By looking at Table 3, we can notice that the two-components specification is characterized by an extremely high entropy level, but a very poor fit. This model only separates the excess of zeros from a more suitable distribution of the outcome. The three-components specification continues to identify a 25% of the municipalities in which we never observed an attack, while the remaining classes contain both zeros and positive attacks. In a four-components model, the size of those components characterized by a positive number of attacks remain stably around 40% and 30%, respectively. Nonetheless, when $K=4$, we identify another fraction of zeros affecting the initial distribution (component 4). The same is true when we consider $K=5$, but the entropy index is smaller than the one obtained with $K=4$. Given these considerations, we opted for a four-components FE-FMM. We present the results for the components characterized by a positive number of attacks.

Table 3: Information criteria and classification results

	2	3	4	5
AIC	-10352.7	-12236.5	-12763.9	-13028.3
BIC	-10162	-11947.2	-12375.9	-12541.6
BIC (Adjusted)	-10254.2	-12087	-12563.4	-12776.8
Entropy	0.976	0.793	0.801	0.793
# free	29	44	59	74
Log likelihood	5205.363	6162.261	6440.934	6588.135
Components				
1	73.24	41.2	40.32	39.37
2	26.76*	33.08	30.64	26.89

3	25.72*	23.95*	23.51*
4		5.09*	5.34*
5			4.88*

* Classes composed by censored observations.

Figure 1 provides a graphical representation of the FE-FMM estimates. To facilitate the comparison, we have also included the FE results presented in Table 2. The values used to construct Figure 1 can be found in the appendix. In general, FE results are driven by Component 1, which represents the largest component (40.32% of observations). Concerning the sensitivity to paramilitaries' attacks, both guerrillas' subgroups have more or less the same behavior: when paramilitaries increase their number of attacks per area by a one percentage point, guerrillas respond with a 2% increase in the offenses. In this respect, Component 2 reacts slightly less than Component 1. Component 2 fights much less in coca production areas. In contrast, Component 1 is more likely to attack northern areas. Therefore, we expect a higher concentration of Component 1 in the Caribbean region. Component 1 tends to attack richer areas. Income distribution does not affect the number of attacks undertaken by the two components. Interestingly, Component 1 is more interested in coal and metal areas, while Component 2 attacks especially oil and coffee areas. This is consistent with the idea of a rural component of guerrillas. Finally, there is a negative correlation between the size of the population and the number of attacks undertaken by Component 2. On average, the number of attacks from Component 1 is 1.15, while the average number of attacks undertaken by Component 2 is 0.67.

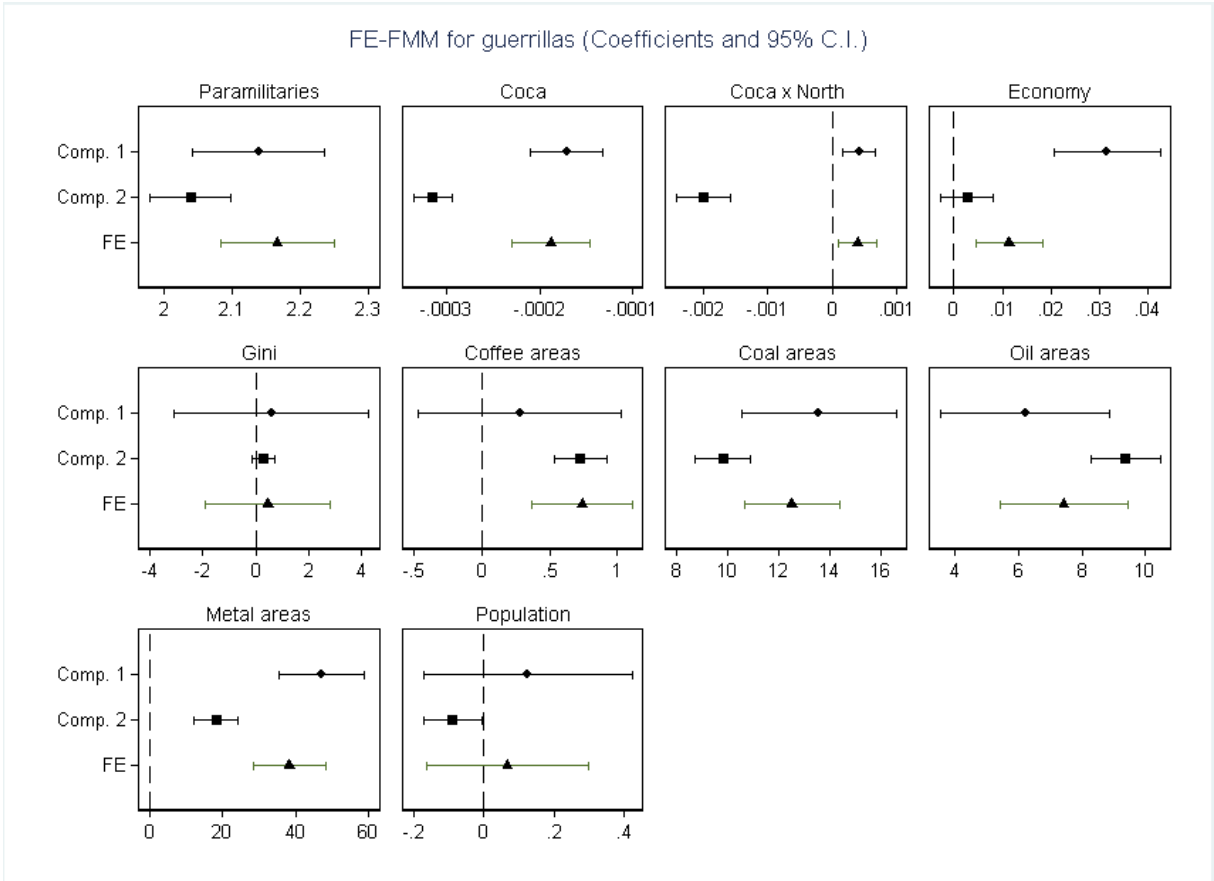


Fig. 1: FE-FMM estimates for guerrillas (coefficients and 95% CI)

Table 4 reports the estimated coefficients of those variables for which a standard FE model leads to extremely biased results. These variables capture the main dimensions of predatory activities and are rather important in the literature on economics of conflict. For instance, if we consider coca production areas, using a linear FE model we would conclude that *guerrilleros* are interested only in those production areas located in the North of the country. In particular, we would expect 4 attacks per ten square km every 10,000 hectares of coca leaf in the North. However, this evidence is consistent only with component 1's behavior, whereas component 2 tends to attack less coca production areas independently of whether they are located in the North or in the South of the country. Similarly, the positive correlation that we have in the first column of Table 4 between the economy of the area and the number of attacks is exclusively due to the component 1's behavior. According to FE estimates, a one standard deviation increase in the economy index is associated with an increase in the number of attacks of 0.022 attacks per ten square kilometers. In fact, once we account for the intra-group heterogeneity, the coefficient of economy for component 1 becomes three times larger than the linear FE estimates,

that is, 0.063 attacks per ten square kilometers every one standard deviation of the economy variable. The same conclusion holds for areas involved in metal extraction activities, where a variation of one standard deviation of the independent variable is associated with an increase in the number of attacks per ten square km of 0.5 when we use a linear FE estimator. This estimate is the weighted average of two different coefficients: 0.6 attacks per ten square km for component 1 and 0.23 attacks per ten square km for component 2. Again, component 1 shows a more predatory behavior than component 2. Finally, with a decrease of 0.08 attacks per ten square km every 10,000 inhabitants, only component 2 is sensitive to population density. This means that component 2 attacks less densely populated areas. Therefore, even if the two components do not differ so much in terms of incidence of attacks, the second component is less sensitive to traditional determinants of conflict. If we do not take into account these differences, our estimates could lead to approximate or even incorrect conclusions and policy recommendations.

Table 4: FE and FE-FMM estimates for guerrilla

	FE	Comp. 1	Comp. 2
	(1)	(2)	(3)
Coca	-0.0002*** (8.651)	-0.0002*** (8.595)	-0.0003*** (30.050)
Coca × North	0.0004** (2.516)	0.0004*** (3.328)	-0.0020*** (9.465)
Economy	0.0114*** (3.291)	0.0316*** (5.657)	0.0028 (0.993)
Metal	38.2493*** (7.627)	46.9271*** (7.878)	18.0940*** (5.756)
Population	0.0675 (0.575)	0.1263 (0.830)	-0.0877** (2.095)
Constant	-8.3617*** (8.042)	0.0042** (2.471)	0.0017 (0.598)
Additional controls	Yes	Yes	Yes
Log-lik	-524.029	6440.934	
Observations	5,303	5,303	5,303
Robust z-statistics in parentheses.			
Significant at: *10%, **5%, ***1%.			

Figure 2 shows municipalities in different color according to the type of attack that is more likely. Recall that a given municipality might experience both types of attacks. It is clear that one of the two types predominates in most municipalities. Municipalities in blue (component 1) are places in

which the war is basically explained for fighting for rents both legal (local government revenues, mining) or illegal (coca). Municipalities in green (component 2) are places that tend to be less urbanized (main cities of the country are marked in blue). This type of attacks are exacerbated by the presence of coffee and oil and placated by the presence of coca crops. This might suggest that coca crops in green municipalities are likely to be under the guerrilla control.

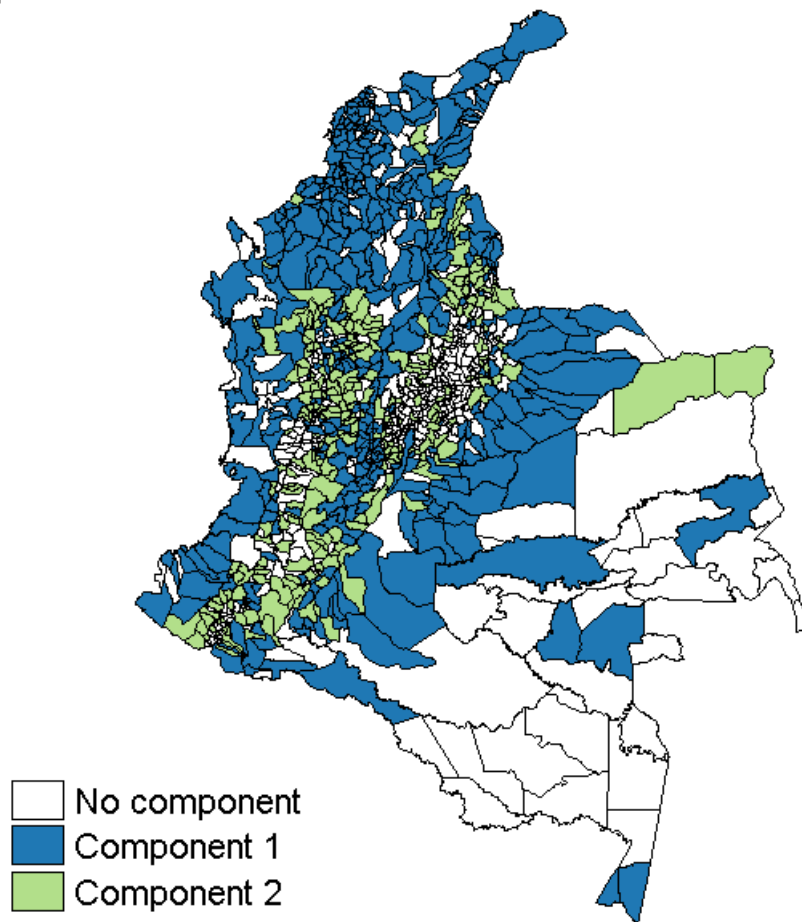


Figure 2: Location of the two components.

Component 1 is very significant in the Caribbean region (departments of Córdoba, Sucre, Atlántico, Bolívar, Magdalena and Cesar) and in the Llanos orientales (departments of Arauca, Casanare and Meta). Both regions are very important for smuggling of illegal drugs and mining activities. Interestingly, Component 2 mainly occupies the Andean región, which is the

most developed in the country. This type of attacks is particularly important in the departments of Huila, Caldas and Quindio.

The main FARC structures (the blocs) are geographical structures that are located in more or less the same areas through time. All blocs cover areas with both types of components except for the central bloc (CE in Figure 3) that is based on the three departments mentioned above. Figure 3 depicts component 2 by FARC blocs. It is very interesting that the areas out of the Andean region correspond to arguably the richest blocs, the eastern bloc (EA, orange), the south bloc (SO, red) and the north-west bloc (NW, light green). Those areas are much correlated to coca production.

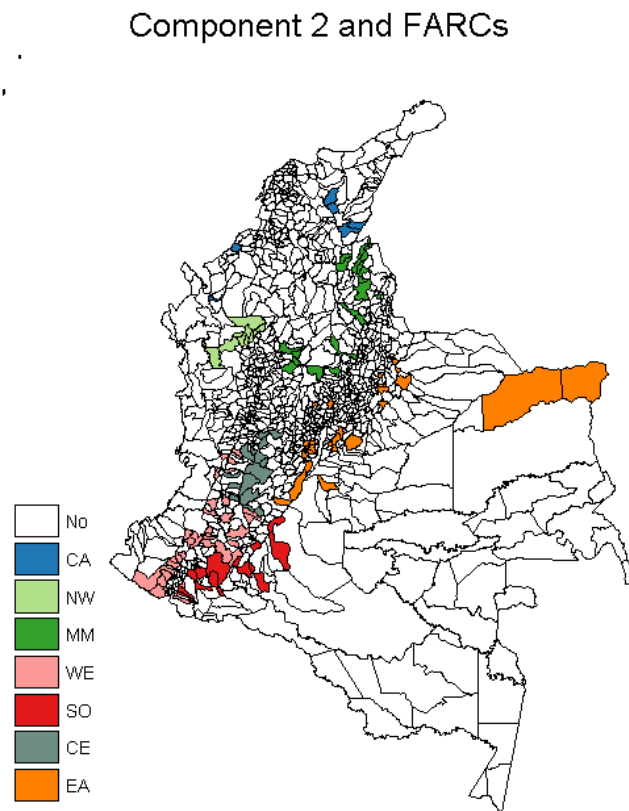


Figure 3. Component 2 by FARC blocs. CA: Caribbean bloc; NW: North-West bloc; MM: Magdalena medio bloc; WE: Western bloc; SO: South bloc; CE: Central bloc; EA: Eastern bloc.

In sum, our technique has distinguished two groups of municipalities, one that is attacked more intensively (C1) and another less intensively (C2). Attacks of C1 type seems to be more likely in areas which the war for grabbing (legal and illegal) rents is strong. Attacks of C2 type are compatible with two different stories: they have to do either with municipalities with coca businesses (production or trafficking) that are presumably under control of FARC or with less urbanized (less developed)

municipalities in the Andean region, which happens to be the most developed region in the country (an easier prey in hard zone).¹²

As a robustness check we have repeated our analysis considering only the period in which President Alvaro Uribe was in power (2002-2005). In 2003, the government approved the Plan Patriota, a policy aiming to affirm state control. Although the number of attacks coming from both components decreased during this period, our conclusions as well as the size of the subgroups remain qualitatively the same. We have also run a FE-FMM using as a dependent variable the number per area of clashes, but in this case no latent class emerged. Finally, to test whether our components capture the presence of ELN in the Colombian conflict, we exploited information on the number of attacks undertaken by ELN in which at least one soldier of ELN was killed. This variable underestimates the true involvement of ELN in the conflict. But, assuming the same probability to be killed in areas belonging to Component 1 and Component 2 (the intensity of the conflict with Paramilitaries and Official Army is not statistically different), we can say that at least 7.3% of Component's 1 areas are characterized by the presence of ELN, while this percentage falls to 3.9% for Component's 2 areas. Overall, ELN faced at least one casualty in 47 out of 1120 municipalities. This means that our results are more likely to be driven by intragroup heterogeneity than by intergroup heterogeneity. Not taking into account intragroup heterogeneity make econometric estimates more likely to be biased.

5. Conclusions

The results of this work are twofold. The first is the novel application of Finite Mixture Models to study behavioral heterogeneity in actual conflicts. Secondly, we have empirically investigated the existence of behavioral differences in Colombian civil conflict in the period 1999-2005. That is, we have highlighted a behavioral heterogeneity in guerrilla armed groups and their distinct economic correlates so showing that different groups do exist and exhibit different patterns of behavior. In particular, whereas

¹² Our measures of institutions are swept out from our estimation by the fixed effects. Simple correlations show that C2 attacks are negatively correlated with all our institutional measures. This is consistent with previous work on the involvement of FARC in several aspects of coca businesses in the south of the country and Llanos orientales (Pardo (2000), Pizarro (2011)); and posterior evolution of the conflict that shows that FARC retreated from the center of the country (Corporación Nuevo Arcoiris, 2013) .

paramilitaries behave as a rather homogenous group, guerrilla attacks can be divided into two different types. Type 1 (C1 throughout the text) is more likely in a large number of municipalities and those municipalities are attacked more oftentimes. Type 2 (C2) is more likely in a smaller number of municipalities and those municipalities are attacked less intensively. Type 1 reflects predatory behavior and type 2 defensive or tactic behavior. Type 1 attacks increase with the presence of both legal and illegal rents. Type 2 decreases with coca crops (defensive behavior) and is concentrated in less urbanized municipalities (with weaker institutions) in the Andean region (the most developed in the country) (tactic behavior). This enriches the existing literature by highlighting that only one component of *guerrilleros* is actually committed to exploitation of rents.

In sum, if we look at the methodological aspect of this work, it might be argued that the use of FMM models in empirical literature on conflict needs further research and additional applications. This paper is nothing but an interesting start in this respect. When taking into account the contribution to the understanding of Colombian conflict, this paper has to be mentioned because it highlighted the heterogeneity of preferences within guerrilla groups which have been often treated in the literature as an indistinct and unitary actor.

References

- Acemoglu, D., Robinson, J. A. and Santos, R. J. (2013) The Monopoly of Violence: Evidence from Colombia. *Journal of the European Economic Association*, 11(S1), 5-44.
- Akcinaroglu, S. (2012) Rebel Interdependencies and Civil War Outcomes. *Journal of Conflict Resolution*, 56(5): 879-903.
- Angrist, J. D. and Kugler A.D. (2008) Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia. *The Review of Economics and Statistics*, 90(2): 191-215.
- Bago d'Uva. T. (2005) Latent Class Models for Use of Primary Care: Evidence From a British Panel. *Health Economics*, 14(3): 873– 892.
- Barrera, F., Ibáñez A., 2004. Does Violence Reduce Investment In Education? A Theoretical And Empirical Approach. Documentos CEDE No. 2004-27. Bogotá: Universidad de los Andes.
- Bloom, M. (2004) Palestinian suicide bombing: public support, market share and outbidding. *Political Science Quarterly*, 119(1): 61-88.
- Bove, V. and Smith, R. (2011). The Economics of Peacekeeping, in: D.L. Braddon and Hartley, K. (eds.) *Handbook on the Economics of Conflict*, Edward Elgar, Cheltenham.
- Brauer, J., Vargas, J.F., (2012) Colombia: Introduction to a Special Issue. *Defence and Peace Economics*, 23(2): 107-108.
- Butler C., Cunningham, D.E. and Gates, S. (2014). Modelling Multi-Party contests: from Intimate alliances to Free-for-Alls, mimeo.
- Cameron, C. and Trivedi, P.K. (2013). *Regression Analysis of Count Data*. Cambridge University Press (2nd edition).
- Caruso, R. and Schneider, F. (2013). Brutality of Jihadist Terrorism. A contest theory perspective and empirical evidence in the period 2002-2010. *Journal of Policy Modeling*, 35(5): 685-696.
- Caruso, R. and Vargas J.F. (2014). Conflict, Crime and Violence in Colombia, Introduction to the special issue, Conflict, Crime and Violence in Colombia. *Peace Economics, Peace Science and Public Policy*, 20(1): 1-4.
- Corporación Nuevo Arcoiris (2013). Informe del observatorio del conflicto colombiano. Bogotá
- Celeux, G. and Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13(2), 195-212.

- Cortés, D., Vargas, J.F., Hincapié, L., Franco, M., (2012). Democratic Security, Police Presence and Conflict in Colombia (In Spanish). *Desarrollo y Sociedad*, 69 (1), 11-32.
- Cortés, D. and Montolio, D. (2014). Provision of Public Goods and Violent Conflict: Evidence from Colombia. *Peace Economics, Peace Science and Public Policy*, 20(1), 143-167.
- Cubel, M. and Sanchez-Pages, S. (2012). The effect of within-group inequality in a conflict against a unitary threat. *Peace Economics, Peace Science and Public Policy*, 18(3), 1-11.
- Deb, P. and Trivedi, P.K. (2013). Finite Mixture for Panels with Fixed Effects. *Journal of Econometric Methods*, 2(1), 35–51.
- Della Porta, D. (1995). *Social movements, political violence and the state*, Cambridge University Press, Cambridge.
- Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1), 1-38.
- Díaz, A.M., Sánchez, F., 2004. *Geography of Illicit Crops (Coca Leaf) and Armed Conflict in Colombia*. The Development Research Centre, Development Studies Institute, London School of Economics (February 2004).
- Dube, O. and Vargas, J. F. (2013) Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *Review of Economic Studies*, 80(4): 1384-1421.
- Esteban, J. and Ray, D. (2011). A Model of Ethnic Conflict. *Journal of the European Economic Association*, 9(3), 496-521.
- Gallego, J. (2011). Civil conflict and voting behavior: Evidence from Colombia. SSRN Working Paper, 1911983.
- Gourieroux, C., Monfort, A. and Trognon, A. (1984) Pseudo maximum likelihood methods: Theory. *Econometrica*, 52(3): 681-700.
- Gupta, D.K. and Mundra, K. (2005) Suicide bombing as a strategic weapon: an Empirical investigation of Hamas and Islamic Jihad. *Terrorism and Political Violence*, 17(4): 573-598.
- Gurr, T. R., (1968) Psychological factors in civil violence. *World Politics*, 20(2): 245-278.
- Hirshleifer, J. (1991) The Paradox of Power. *Economics and Politics*, 3(3): 177-200.

- Idrobo, N., Mejía, D. and Tribin, A. (2014) Illegal Gold Mining and Violence in Colombia. *Peace Economics, Peace Science and Public Policy*, 20(1): pp. 83-111.
- Lemus, N. (2014). Conflict-Induced Poverty: Evidence from Colombia. *Peace Economics, Peace Science and Public Policy*, 20(1): 113-142.
- Lubke, G. and Muthén, B. (2007). Performance of factor mixture models as a function of model size, covariates effects, and class-specific parameters. *Structural Equation Modeling*, 14(1), 26-47.
- Medina, C. (2010). FARC-EP and ELN. A comparative political history (1958-2006) (In Spanish). Ph. D Thesis. National University of Colombia.
- Mingorance, F., Minelli, F. and Du, H. L. (2004). El cultivo de la palma africana en el Choco. Human Rights Everywhere, HREV.
- Moor, M. and van de Sandt, J. (2014) The Dark Side of Coal: Paramilitary Violence in the Mining Region of Cesar, Colombia. PAX, Utrecht, The Netherlands.
- Nilsson, D. (2008) Partial Peace: Rebel Groups Inside and Outside of Civil War Settlements. *Journal of Peace Research*, 45(4): 479-495.
- Nilsson, D. (2010) Turning Weakness into Strength: Military Capabilities, Multiple Rebel Groups and Negotiated Settlements. *Conflict Management and Peace Science*, 27(3): 253-271.
- Nussio, E. (2011) Learning from Shortcomings: The Demobilisation of Paramilitaries in Colombia. *Journal of Peacebuilding & Development*, 6(2): 88-92.
- Offstein, N. (2003) An Historical Review and Analysis of Colombian Guerilla Movements. *Revista Desarrollo y Sociedad*, 52(1), 99-142.
- Palacios, P. (2012). Forced Displacement: Legal Versus Illegal Crops. *Defence and Peace Economics*, 23 (2): 133-160.
- Pardo, R. (2000). Colombia's two-front war. *Foreign Affairs*, 79 (4): 64-73.
- Pizarro, E. (2011). Las FARC (1949-2011): De guerrilla campesina a máquina de Guerra. Grupo Editorial Norma, Bogotá.
- Restrepo, J., Spagat, M. and Vargas, J. F. (2004). The Dynamics of the Colombian Civil Conflict: A New Dataset. *Homo Oeconomicus*, 21(3): 396-429.
- Rodriguez, M., Daza, N., (2012) Determinants of Civil conflict in Colombia: How Robust are They? *Defence and Peace Economics*, 23 (2): 109-131.
- Rodríguez, C., Sánchez, F., 2009. Armed Conflict Exposure, Human Capital Investments and Child Labor: Evidence from Colombia. Documentos CEDE No. 2009-05. Bogotá: Universidad de los Andes.

- Rodríguez, C., Sánchez, F., 2010. Books and Guns. The Quality of Schools in Conflict Zones. Documentos CEDE No. 2010-38. Bogotá: Universidad de los Andes.
- Sanchez, F. and Palau, M. (2006), Conflict, Decentralization and Local Governance in Colombia, 1974-2004. (CEDEWorking Paper No. 2006-20).
- Skrondal, A., and S. Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling*. New York: Chapman & Hall.
- Vargas, J. F. (2012) The persistent Colombian conflict: subnational analysis of the duration of violence. *Defence and Peace Economics*, 23(2): 203-223.

Appendix A: FE-FMM Coefficients

Table A1 reports the coefficients we use to construct Figure 1.

	FE	Comp. 1	Comp. 2
	(1)	(2)	(3)
Rival group	2.1661*** (50.670)	2.1388*** (43.314)	2.0387*** (66.986)
FS residual	-1.9846*** (341.560)	-1.9667*** (232.349)	-2.0272*** (251.190)
Coca	-0.0002*** (8.651)	-0.0002*** (8.595)	-0.0003*** (30.050)
Coca × North	0.0004** (2.516)	0.0004*** (3.328)	-0.0020*** (9.465)
Economy	0.0114*** (3.291)	0.0316*** (5.657)	0.0028 (0.993)
Gini	0.4701 (0.392)	0.5873 (0.312)	0.3053 (1.377)
Coffee	0.7410*** (3.917)	0.2820 (0.738)	0.7279*** (7.517)
Coal	12.5132*** (13.059)	13.5840*** (8.758)	9.8032*** (17.815)
Oil	7.4337*** (7.264)	6.2201*** (4.605)	9.3733*** (16.961)
Metal	38.2493*** (7.627)	46.9271*** (7.878)	18.0940*** (5.756)
Population	0.0675 (0.575)	0.1263 (0.830)	-0.0877** (2.095)
Mills	1.8205*** (4.287)	1.8185*** (3.627)	0.0767 (0.271)
Constant	-8.3617*** (8.042)	0.0042** (2.471)	0.0017 (0.598)
Log-lik	-524.029	6440.934	
Observations	5,303	5,303	5,303

Robust z-statistics in parentheses.
Significant at: *10%, **5%, ***1%.