Conflict and roads in Colombia

Juan Felipe González
Msc in economics, Universitat de Barcelona

Advisors: Enrique López-Bazo♣ & Daniel Albalate del Sol♠
♣ Universitat de Barcelona (AQR-IREA)
♠ Universitat de Barcelona (GiM)

Abstract

The civil unrest in Colombia has reported over 8 million victims. This paper studies the effect of transport infrastructure on the level of conflict at the municipality level. I confirm the main hypothesis that roads have a significant and negative effect on conflict, that is, better roads endowment discourage conflict in general. I also find that populated municipalities are the primary shelter for displaced. However, the agglomeration effects do not overshadow the road's effect. I interpret these results as inter-regional roads increase the state capacity and shift the economic activity in the regions, the latter decreases the probability of participating in the conflict by either joining armed groups or being a victim. The role of roads is found to be robust to a long list of controls, different specifications, different estimation techniques, and potential bias from simultaneity and measurement error.

Keywords: Transport infrastructure, roads, conflict, displacement, population density, regional economics, Colombia.

ACKNOWLEDGEMENTS: First, I would like to express my gratitude to my advisors Enrique López-Bazo and Daniel Albalate. Thanks to QQ for his exceptional guidance, expertise, feedbacks and to make myself question about every single detail. Thanks to Daniel for his helpful explanations, knowledge, and keep me motivated throughout this process. I am grateful to Gilles Duranton for providing me with the data. Financial support from the Spanish Ministry of Education and UB school of economics is gratefully acknowledged to achieve this master successfully. This paper is dedicated to Maria Fernanda Cortes (my mom). No-one has contributed more than her to complete this process through her sacrifice, resilience, and love. To Maria Alejandra Suárez to her unconditional support and make this stage of my life wonderful. Thanks to my friends Gaby and Vincent for their advise and make my life easier and enjoyable during these two years. Last but not least, thanks to all the team of UB School of Economics, especially Jordi Roca for his incredible effort and commitment with us.

1 Introduction

Colombia has suffered one of the most protracted civil war in the world which has had several consequences concerning victims. The victims unit review (RUV) has reported more than 8 million victims due to the armed conflict which the 73% of them are because of forced displacement. This fact makes Colombia the country with one of the largest population displaced in the world. The more critical period was from 1995 to 2015. In 2016, the Colombian government and the insurgent group Revolutionary Armed Forces of Colombia (by its acronym in Spanish FARC) signed a peace agreement, this meant a considerable step forward from both parts.

Simultaneously, in 2014 the government started one of the most ambitious projects in transport infrastructure. The project plans to build and repair more than 8.000 km of roads with an estimated investment of $US18.000 millions what represents around the 6-7% of the Colombian GDP. The primary objective of the project is to improve competitiveness, get access to the periphery of the country and increase the presence of the state in those regions.

The literature on conflict has focused on trying to explain the causes of the conflict and the literature on transport infrastructure has concentrated mainly on the impact on the economic outcomes which is ambiguous (Melo et al. (2012)). The direct relationship between conflict and transport infrastructure has not been explored yet in the literature. Thus, this paper attempts to fill this gap in the literature.

All of this brings up to different questions. First, does better transport infrastructure contributes to reducing conflict? Second, if it does, in which magnitude does it contribute? Also, which might be the possible mechanisms behind?

To make progress on this research agenda, this paper investigates the effects of roads on forced displacement for Colombian municipalities. I first confirm that higher density roads contribute substantially to reduce forced displacement. In particular, increasing by double the stock of roads in a municipality the amount of displaced decreases between 3-7%. Estimates using instrumental variables suggest even stronger effects of 11% reduction of conflict when the stock of roads increases by double. I interpret these results as roads increase the presence and legitimacy of the Colombian government over the territory, Walder (1995); Besley and Persson (2010) substantiate these results. Moreover, roads incentive regional development, especially impacting economic outcomes as trade, productivity, firms location and diversification of the economy and participation of the formal labor market (Duranton (2015); Cárdenas and Sandoval (2008); Cárdenas (2002)). All of this contributes to decrease the incentive to participate in the conflict (either as insurgent or victim).

Studying the link between conflict and transport infrastructure for Colombia is appealing from different perspectives. Colombia is a mid-income country where despite the

conflict, it is a country which has been growing around 4% per year (GDP) during the last fifteen years. Indeed, due to the great performance during the last years, Colombia was added as a member of The Organisation for Economic Co-operation and Development (OECD). The country also has harsh geography, divided by the high mountains in the Andean region, and the extensive lands of the rainforest. Besides the harsh geography, the negligence in previous years of the Colombian government to develop the road network might have helped to propagate the conflict. Hence, this paper especially matters for Colombian public policies, especially to move forward and remove the stone on the way which is civil unrest. This paper confirms the necessity of the 4G roads project development. Additionally, investigating transport infrastructure in a developing country as Colombia contributes to extending the knowledge in the field since very little is known about this kind of countries and especially its impact on social outcomes as conflict.

The rest of this paper is organized as follows. Section 2 describes the Colombian context. Section 3 discusses previous studies on the field. The description of the data is detailed in section 4. The empirical model and the results are presented in section 5 and 6. Section 7 discusses the results, and finally section 8 concludes.

2 Context

Colombia has faced different types of social struggles throughout its history which are tied to the colonial period. This study focuses on the current conflict which started with the launch of a communist insurgency in the decades of the 60's motivated by the victory of the Cuban revolution and the consequences of the bipartisan war or the period so-called "La Violencia" (1948 - 1958).

In some of the essays developed by the National Center of Historical Memory (CNMH by its Spanish acronym) are presented different reasons of the origins of the conflict. For instance, Wills (2015) points out the particularities of the formation of the Colombian Nation-State. In fact, Kalmanovitz (2003) describes the society as fragmented, with a very sick ideology of national unity, whose different interests can only find a solution, and with difficulty, in a State that is also divided. He also highlights the heterogeneity of the economic structure which persists nowadays.

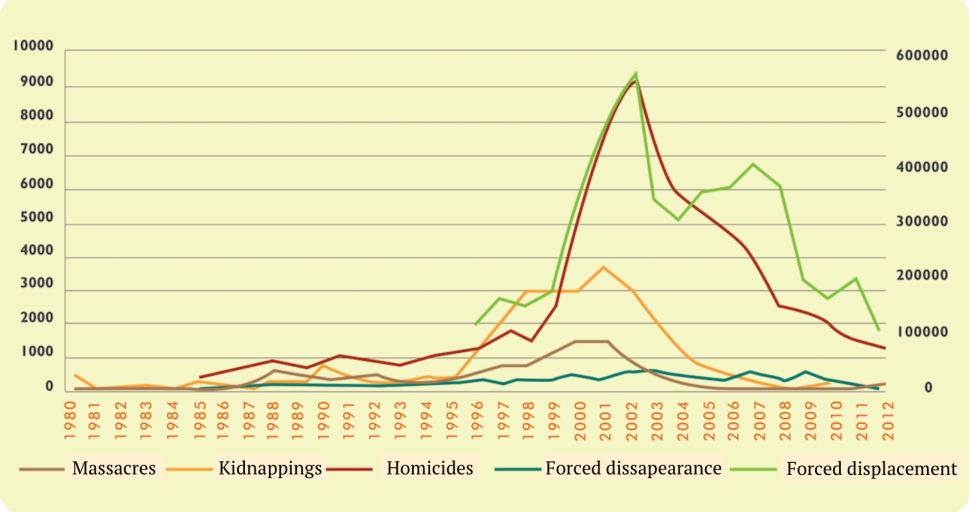
Fajardo Montaña (2015) offers a different perspective, he states that the origins and the main problem are due to the distribution, appropriation and the use of the land. First, the agricultural model was focused on the settled in the large property, then, moving on the profits of the drug trafficking business and nowadays moving back to the benefit of the large landowners. The influential groups have used different appropriation and control mechanisms, especially separating the communities from their lands and redistributing among their groups.

The combination of those mentioned above might explain other consequences. For

instance, the primary needs unsatisfied is a consequence of the land inequality and the absence of the nation-State across the country, especially in the periphery. The conflict has taken place in the rural area of the country, out of the so-called gold triangle [1] except for the narco-traffickers war during the 90's. Even though the conflict has taken place in the periphery, the core has also suffered the consequences through displacement, and this has had implications for the development of the cities. Conforming to the different censuses, by 1938, the rural population was about 71%, by 1964 around 48%, and between the 90's and 2016 around 30% and 20% respectively. At the end of the year 2017, Colombia had a population of about 49 million over a territory of 1.1 million km^2. The most of the people live either along the Caribbean coast or on the high plateaus of the interior.

During the conflict four set of actors have been involved: left-wing guerrillas[2], the government, the right-wing paramilitaries[3], and organized criminal groups (BACRIM and drug cartels). The conflict has been present since its origins, however, its intensity has varied over time. As reported by CNMH (2013) the conflict remained low intensity in the 1980s, but it started to rush to high intensity at the beginning of the 1990s, reaching its peak in 2002 (see figure 1). The increase in conflict might be due to the presence of paramilitary groups and the control of the drug business under the guerrillas' groups.

Figure 1: Evolution of the Colombian conflict



Source: Grupo de memoria Historica (GMH)

Paramilitary groups were created to re-take the control of different territories occupied by the guerrillas in response to guerrillas' extortion. Rural landowners and drug barons formed these groups, and they operated in several times jointly with the Colombian army. The war between the guerrillas and paramilitary groups took place in the periphery of

[1] It is formed by the three main cities of the country (Bogotá, Cali, and Medellín).

[2] In this group, one may include the Armed Revolutionary Forces of Colombia (by its Spanish acronym FARC), Ejercito de liberación Nacional (ELN), Movimiento 19 de Abril (M-19), among others.

[3] Autodefensas Unidas de Colombia (AUC), Los pepes and CONVIVIR).

the country where the group settled in the region was the ruler, dictated its constitution, justice and regulated the economy. The presence of insurgent groups blocked any attempt of the government to intervene in the region, being isolated from the center allowed them to control the cocaine business, steal natural resources, and commit crimes.

The report of CNMH (2013) defines the Colombian conflict as frequent and low intense which had a little impact at the national level, but with a high incidence at the local level. Despite this, the center remained growing and developing, at least in the last 30 years the country has shown positive GDP growth rate (except in the year 2009). These two faces of the nation just increased the inequality gap between core and periphery.

After 60 years of conflict, the government and the most significant insurgent group FARC signed a peace agreement in September 2016. Previously in 1998, the government failed in their attempt[4] to reach a deal. Under the administration of Uribe from 2003 to 2006, a demobilization process was carried out with the paramilitary groups. However, it presented different failures that lead their members to get back in the conflict through organized criminal groups. The peace agreement signed with the FARC was a huge step since it corrected the mistakes done in previous negotiations. The negotiations took place in Havana during four years where Cuba and Norway acted as guarantee parts in the process. To reach the agreement, the society, the government and the FARC debated the six points[5] proposed in the initial negotiations under a ceasefire environment previously agreed.

The peace agreement came along with the most ambitious project ever of transport infrastructure in Colombia called fourth generation (4G) - roads of prosperity. According to the National Infrastructure Agency (ANI), this project aims to build and operate over 8000 km of roads with an investment of 47 billion COP[6]. The primary objectives of the project are to improve the competitiveness, get access to isolated regions, connect the core-periphery and finally settle the presence of the state in the forgotten areas in the previous decades.

A set of different factors might explain Colombian conflict. This paper attempts to identify and quantify the role of transport infrastructure and agglomeration in cities through empirical evidence. The level of roads deployment is a suitable mechanism for the government to show presence over territory and bring development to the region.

In the decade of 30's, the government in charge decided to change the system of railroads by the system of roads which would conduct the first roads project in the country. During the period 1950 - 1958 the second wave of projects were carried out. Also during 1975 and 1994, the national road network increased by 5.600 km, and the regional roads increased its length in 55.000 km. After the early 90's the government has given priority

[4] Caguan negotiations.
[5] See the peace agreement for more information Presidencia de la Republica de Colombia (2016).
[6] Around 18000 millions of USD.

to other public policies as education and health. The lack of interest during the last two decades contributes to the infrastructure development backwardness. Therefore, the country lost competitiveness and increased regional disparities. Concerning roads, by 2010, the country had large regional disparities and it was not well connected between regions.

Colombia is facing a crucial moment in its history, either it overcomes its problems and reach a long and lasting peace or remains in the scenario of conflict. Achieving the first one might encourage a faster agreement with the ELN[7].

3 Literature review

The relationship between conflict and transport infrastructure is difficult to establish and especially to measure it, since it has not been studied extensively as a direct impact. However, there are some political and economic mechanisms behind that may capture the impact of roads on the civil conflict.

The primary element behind is the state capacity. Evans et al. (1985) define the state capacity as the ability of a government to administer its territory effectively. In this context, Walder (1995) highlights four basic state capacities in the modern world: 'the capacity to mobilize financial resources from the society to pursue what the central policymakers perceive as the "national interest" (extractive capacity); the capacity to guide national socio-economic development (steering capacity); the capacity to dominate by using symbols and creating consensus (legitimation capacity); and the capacity to dominate by the use or threat of force (coercive capacity)'.

A first mechanism is related with the legitimation and coercive capacity. The governments develop their road network that facilitates the fast movement of troops to be ready for internal repression and external intervention. In particular, Saiz (2006) explores this idea from the dictatorships perspective. Following Pérez (2005) during the years 1950 and 1958 the development of roads increased rapidly in Colombia, even though the average annual growth in roads overcame the GDP growth rate. During this period[8], the government repaired and built around 4600 km of roads. This idea has its roots in the ancient Rome where more than 55,000 thousands miles of paved highways were built. As Hitchner (2012) states, "it is important to recognize that the primary purpose of Roman roads was for state and military functions, this load- and wheel-bearing standard of construction allowed it to transcend this limited purpose."

A second mechanism is linked with the steering capacity. Development in public infrastructure supposes to bring economic benefits to the different regions. Similarly, regions

[7]In February of 2017, the government began talks with the ELN.

[8]This period coincides with the unique dictatorship led by Rojas Pinilla in Colombia during the XX century.

with better socio-economic conditions should have a lower level of conflicts. Particularly, Besley and Persson (2010) present a theoretical model associating weak states with poor economic performance. They show that low legal capacity can be conducive to lackluster economic growth or might contribute to the likelihood of civil war.

The paper by Fearon and Laitin (2003) supports this idea with an empirical analysis using a panel at a country[9] level. They find that the prevalence of internal war is mainly the result of a steady accumulation of protracted conflicts since the 50s and 60s what is in line with the roots of the Colombian conflict. Additionally, the authors suggest that the factors that explain civil war are conditions that favor insurgency, such as poverty, slow growth, weak states, rough terrain and large populations.

Testing the greed and grievance theories Collier and Hoefflery (2004) find similar results, greed[10] theory outperforms grievance[11]. According to their conclusions, greater ethnic and religious diversity reduce the risk of conflict. On the contrary, both dependence upon primary commodity exports and a large diaspora substantially increase the risk of conflict. The latter might be conditions that favor insurgency. Indeed, Colombia is a particular case of primary commodity export dependence along its history, and it has suffered one of the largest 'internal - diaspora' during the last 60 years.

In this perspective, transport infrastructure might be a road hump on the conflict through some economic mechanisms. However, taking a closer view on the infrastructure literature and in particular transport infrastructure, one may notice that the debate about whether or not infrastructure brings development and growth to the regions is still open.

In the late 80's, Aschauer (1989) provided the basis for the discussion. He finds that the core infrastructure[12] is an essential factor to explain the productivity in a region. Since core infrastructure is more accessible to measure and represents about the 50% - 60% of the public capital, most of the studies on the topic focus in this type of infrastructures.

At the early 90s, the Aschauer's study was hardly criticized, especially on account of econometric flaws. For instance, Gramlich (1994); Holtz-Eakin (1994); Eisner (1991) have rejected Aschauer's results because of causality and spurious correlation. What is more, most of them found that public capital investments led to the widening of regional disparities. Bearing this in mind, subsequent studies as, (Evans and Karras, 1994; Hulten and Schwab, 1991; Holtz-Eakin and Schwartz, 1995; Álvarez et al., 2006) find no significant evidence of the positive effect of public capital on productivity.

[9]They show that even countries with different conflict ideology characteristics as Islamic fundamentalist(Afghanistan, Algeria or Kashmir), right-wing reaction (Nicaragua), or ethnic nationalism (Sri Lanka, Turkey, Rwanda, Burundi, Sudan, Ethiopia, India, Indonesia) and communism in Southeast Asia or Latin America have conducted the conflict through almost entirely in the form of insurgency or rural guerrillas.

[10]The 'greed' model focuses on the sources of finance of civil war.

[11]The 'grievance' model examines inequality, political oppression, and ethnic and religious divisions as causes of conflict.

[12]Core infrastructure is the one which fosters the economic growth and improves productivity (i.e., motorways, railways airports, utilities, mass transit, among others)

Similarly, additional studies find negative effect especially of transport infrastructure on economic outputs. For instance, Crescenzi and Rodríguez-Pose (2012) suggest that transport infrastructure endowment is a poor predictor of economic growth in the European Union. Likewise, Boarnet (1998) shows that negative spillover effects exist in the case of street and highway capital in California. More recently Duranton et al. (2013) explore the causal effect of highways on trade. They find a small and often insignificant effect of roads on trade in value for the case of the United States.

On the contrary, more recent studies which deal with econometrics flaws find positive effects of transport infrastructure on economic outcomes. For instance, Donaldson (2010) investigates the impact of India's rail-roads network, achieving three main results. Firstly, it decreased trade costs and interregional price gaps. Secondly, it increased real income levels and thirdly, it decreased interregional and international trade. Consistent with this, Percoco (2016) obtains that the construction of the Italian roads between the 1950 and 2001 had a positive impact in terms of urban development, employment growth, and firm entry.

Focusing on the case of Colombia, Duranton (2015) finds that the road distance between cities is a significant impediment to trade. He also concludes that major roads crossing a city have a large effect on city's exports and imports. Analysing the role of transport infrastructure in determining plant-level TFP (total factor productivity) Cárdenas and Sandoval (2008) suggests that higher road density has a positive effect on TFP. On top of that, in a previous paper, Cárdenas (2002) finds that high crime is associated with low productivity which is in line with the premise of this paper.

Summing up the literature of transport infrastructure on economic outputs is vast. The literature explores different outcomes such as the economic growth, productivity, labor market, firm's location, trade, regional disparities and among others where there is no agreement among those. The particular conditions (geography, pre-existing infrastructure endowment, policies carried out, social features) of each region and the methodology carried out might conduct different results. Therefore, it is important to take into account these particularities and analyze the results without generalization. In this way, as far as this study is concerned, the link between transport infrastructure and conflict is barely explored.

One of the most striking features is how Colombian cities have been developed. Here, agglomeration process has taken a particular form since a significant part of it might be explained by displacement. Thus, classical economic forces might explain just a little about agglomeration in Colombia. In this context, agglomerated municipalities should present a lower level of conflict.

According to Masahisa Fujita (2003) model agglomeration should lead to positive economic growth and faster innovation. In line with Castells-Quintana and Royuela (2014), the results in the empirical evidence vary depending on the stage of development and

income distribution in the region. Additionally, Castells-Quintana (2017) argues that in developing countries low coverage of basic services is likely to handicap the benefits from agglomeration, increasing congestion costs in terms of transport costs, disease transmission, pollution, conflict, and crime.

Although Colombia is a developing and growing country, displacement extensively explains the agglomeration process. Duranton (2016) finds considerable benefits from agglomeration that may overcome the costs. The author finds an elasticity of wages with respect to population of about 5%. In particular, moving from a city with 10,000 inhabitants to Bogota is associated with about 40 percent higher wages. He also finds that neither urban amenities nor roads affect wages. Based on New Economic Geography studies, Guevara Rosero (2017) examines the effect of regional trade openness on agglomeration in Colombia. Her results suggest that trade enhances spatial agglomeration within regions with a large home market and location advantages. Similarly, Duranton (2015) suggest that the level of roads endowment affects trade and the organization of economic activity, in other words, a large endowment of roads might be seen as a location advantage condition.

This study examines three elements; conflict, transport infrastructure, and agglomeration. The literature has found positive effects on transport infrastructure and agglomerated economies. However, the link between transport infrastructure or agglomeration and conflict is not clear. Researchers have focused on studying the determinants of conflict through direct economic factors, and the lack of this 'economic factors' are present because of enough economic and social incentive to participate in the conflict. Conversely, I attempt to investigate the conditions that might help to decrease the incidence of the conflict. These conditions are a necessary situation to generate the economic incentive to not participate in the conflict. Thus, turning attention to a local level, I attempt to fill the gap in the literature using municipality data of Colombia.

4 Data and descriptive analysis

The performed analysis relies on cross-section data for 1056 municipalities for Colombia, for the year 2010 to study the relationship between civil conflict and transport infrastructure. The data comes from different sources; the primary sources are Centro de Estudios Sobre Desarrollo Económico (CEDE), the data provided by Gilles Duranton (Duranton (2015)), Registro Unico de Victimas (RUV) and Departamento Administrativo Nacional de Estadística (DANE).

The CEDE presents different municipal level indicators such as agriculture, government, general characteristics, education, health, violence and conflict. The primary information is collected from official sources. The data provided by Duranton gathers the main information for transport infrastructure, and the RUV provides the primary information

about the internal armed conflict.

Even though Colombia has reported thousands of deaths during the conflict, displacement is the most representative fact that represents Colombian civil unrest. Also it is easier to identify the status of displaced as fact due to the armed conflict. Other facts (e.g., homicides, kidnapping, physical injuries, threats, etc.) might be more difficult to identify with the armed conflict. Additionally, investigating displacement is interesting concerning regional studies, since the mobility of population due to the civil unrest causes new agglomeration forces. According to RUV, displacement represents about the 75% of the victimizing facts in Colombia. Different authors examine the determinants of displacement, often in their results displacement is significantly explained by facts related to the conflict. According to this, displacement is used as a measurement of conflict[13]. Displacement is defined as the number of displaced over the population[14] of the year 2010. The level of displacement in the year 2010 is lower than the previous years but still significant.

$$Displacement\ proportion_i = \frac{displaced_i}{Population_i}$$

As a robustness check, I use the variable More Victims which is the sum of the population that has been a victim of crimes against sexual integrity, forced disappearance, homicide, kidnapping, personal physical injuries, tortures, and seclusion of Children to armed groups over the population in the municipality. Notice that all these facts may be sum since all of them represents victims and then expressed as a proportion of the population.

In a like manner as Duranton (2015), the present analysis focuses on the road index for the year 2010 [15] as a measurement of transport infrastructure.

$$Road\ Index_i = Log(Freqroad_i) + Log(Exits_i) + Log(mileage_i)$$

According to the annual report of the Ministry of transport for the year 2012, road transport represented around the 73% of the transport services GDP, air transport the 10% and the other means of transport (rail transport, fluvial and maritime transport, and among others) the 16% [16]. Therefore, data related to road infrastructure is the most accurate one to measure transport infrastructure.

There is a frequent association of Agglomerated cities with a better transport infrastructure endowment, for this reason, I include the variable population density to examine

[13]Kondylis (2010) uses the displaced status as conflict measure for Bosnia and Herzegovina. Ibáñez and Vélez (2008) examine the factors that might explain displacement in Colombia. Also, Czaika and Kis-Katos (2009) investigate the determinants of displacement at village level for Indonesia.

[14]The population is estimations by DANE based on the census of 2005.

[15]It is a function of the logs of road frequency (Number of principal roads), number of exits from the municipality by principal roads and road lengths(mileage of these roads).

[16]The participation has not changed much in the following years.

whether the impact of transport infrastructures might be captured by agglomerated municipalities. Population density is defined as the population in the municipality over the land area of the municipality.

Since data is scarce for the road infrastructure, I rely on information for the year 2010. Therefore, this study considers a vast set of control variables (see table 1) to control for factors at municipality level that may affect both transport infrastructure and conflict. The first group of them is the region controls which include a set of dummy variables indicating if the municipality belongs to one of the Colombian macro regions, those are; Caribbean, Pacific, Andean, Amazonian or Orinoquia. The second set is the geographical controls; it includes the altitude, climate, soil aptitude, precipitation, erosion, water, coordinates, and area. The third group is the resources controls which are a set of dummy variables; it controls whether in the municipality there is coal, gold, petroleum, emeralds or cattle. Finally, conflict controls, it is a set of dummies which control whether there is a presence of coca crops and rebel groups.

4.1 Descriptive statistics

Table 1: Descriptive statistics

VARIABLES	mean	sd	min	max
Main variables				
disp_2010	0.00721	0.0124	0	0.108
Displacement	186.1	581.1	0	11,554
road index 2010	3.752	3.596	-6.513	9.740
Pop density	156.6	704.4	0.0939	15,207
More victims	.00054	.0010	0	0.0095
Instruments				
length roads 1938	3.751	8.625	0	81.98
dist roads 1938	38.38	71.64	0	875.4
Pop_dens_1993	130.7	524.6	0.141	11,375
Pop_dens_1985	95.65	392.3	0	8,095
Pop_dens_1964	67.25	208.7	0	4,005
Pop_dens_1951	45.33	94.01	0	1,684
Regional controls				
Andean	0.533	0.499	0	1
Caribbean	0.187	0.390	0	1
Pacific	0.169	0.375	0	1
Orinoquia	0.0559	0.230	0	1
amazonic	0.0559	0.230	0	1
Geographical controls				
Altitude	1,053	876.1	2	3,087
climate	1.926	0.904	1	3
area	946.7	2,655	15	42,178
water	3.352e+06	545,952	0	5.626e+06
aptitude	2.723	1.238	0	8
precipitation	1,952	1,068	160	8,000
erosion	1.961	1.048	0	5
Natural resources controls				
dummy coal	0.0671	0.250	0	1
dummy emeralds	0.00814	0.0899	0	1
dummy gold	0.132	0.339	0	1
dummy oil	0.125	0.331	0	1
dummy cattle	0.658	0.475	0	1
Conflict controls				
COCA 2010	0.197	0.398	0	1
ELN 2010	0.152	0.359	0	1
FARC 2010	0.375	0.484	0	1
AUC 2010	0.0152	0.122	0	1

Notes: Descriptive statistics for 1056 municipalities. Disp_2010 is the dependent variable which is expressed as the proportion of displaced. All the other variables are defined in the appendix (see appendix)

Table 1 presents some descriptive statistics of all variables at the municipal level. Note that, although the main regressions below are estimated for fewer municipalities, these

descriptives statics concern an extensive set of Colombian municipalities. In the second part of table 1 one may observe past populations with values of zero, this because after the constitution of 1991 many of the current municipalities were defined. The mean of the dichotomous variables should be interpreted as the percentage of municipalities with values of one

Figure 2 shows that municipalities in the Amazonian and Pacific were the most affected by displacement in relative terms for the year 2010. Additionally, the departments of Antioquia and Arauca show a high intensity of displacement. The average of displaced was about 186 people per municipality, the municipality with more displaced in absolute terms was Tumaco[17] with 11554 inhabitants displaced, and Nechí [18] was the municipality with more displaced (10.8% of the population) in relative terms.

A significant worry with the variable displacement is that the 20% of the data are zeros. Some of them could be zeros while some other might be missing observations whereby, this variable needs a special treatment[19]. First, I study the zeros carefully by departments. I find that in departments like Amazonas, Boyacá, and Cundinamarca more than 50% of the municipalities are represented by zeros. Also, there are other departments with zeros, but they do not overcome the 50% threshold. For the case of Amazonas is plausible that many municipalities have values of zero, since it is a sparsely populated department. On the contrary, Boyacá and Cundinamarca are one of most populated departments. Although these departments have been receptors of displaced in overall, these have not been spared from the conflict. Indeed, different national report (Nuñez, 2007; Presidencial and Vicepresidencia, 2007) describe the conflict in these regions showing that since 2006 the number of displaced started to increase.

This fact brings up to treat in a particular manner the departments of Cundinamarca and Boyacá. So as not to be unfair, I keep some observations and drop 66 municipalities of the sample. This procedure reduces the proportion of zeros at 14% of the sample. Since I have the 'reception of displaced' index, I assume that all the municipalities that did not have displacement but received displaced should be treated as zeros. On the other hand, those municipalities that did not present displacement and did not receive displaced should be treated as missing observations, and thus exclude them from the analysis.

Figure 3 presents the distribution of the road endowment. One may notice that a large part of the country in 2010 had a poor road infrastructure, especially in the Amazonian and the Pacific regions. Another fact to note is that the country is not well regionally connected. For instance, the Caribbean region (north part of the country) is not well connected with the Andean region, and most of the Amazonian, Pacific, and Orinoquia

[17]It is located in the south-west of the country in the Pacific region.

[18]It is located in the department of Antioquia

[19]The log transformation could be a possible solution to the problem of the problem of the proportion in the dependent variable, however as Burger et al. (2009) explains, the log transformation might create biased and inefficient estimates. See also Santos Silva, J.M.C, Tenreyro (2006); Stewart (2013)

municipalities are disconnected. In fact, many municipalities in these regions have no access by road. Paradoxically, the Andean region is the one with more difficulties[20] to build roads but with better road endowment. This is because the connection between the three main cities is through this region. Something to point out is that the endowment of roads is large around Bogotá, Medellín, and Cali but in municipalities in the halfway between those, roads are scarce.

Colombia is a country with a significant amount of cities with a substantial number of population but still with opportunities for growth. In average the municipalities have a population density (see table 1) of 156.6 per square kilometer. Looking at figure 5 (see appendix), the most populated municipalities are around the main cities[21] and in the coffee region. Moreover, looking at the instrument in table 1, since the 50's the municipalities have been growing regarding population.

Tabel 2 reports the correlation coefficients among the main variables. According to the hypotheses, either the road index and population density are negatively and statistically significantly correlated with displacement. Additionally, population density and road index are statistically significant and positively correlated, that is, more populated municipalities are associated with better road infrastructure as the literature indicates. The variable more victims is positively correlated with displacement and population density but not with the road index.

Table 2: Correlations between the main variables

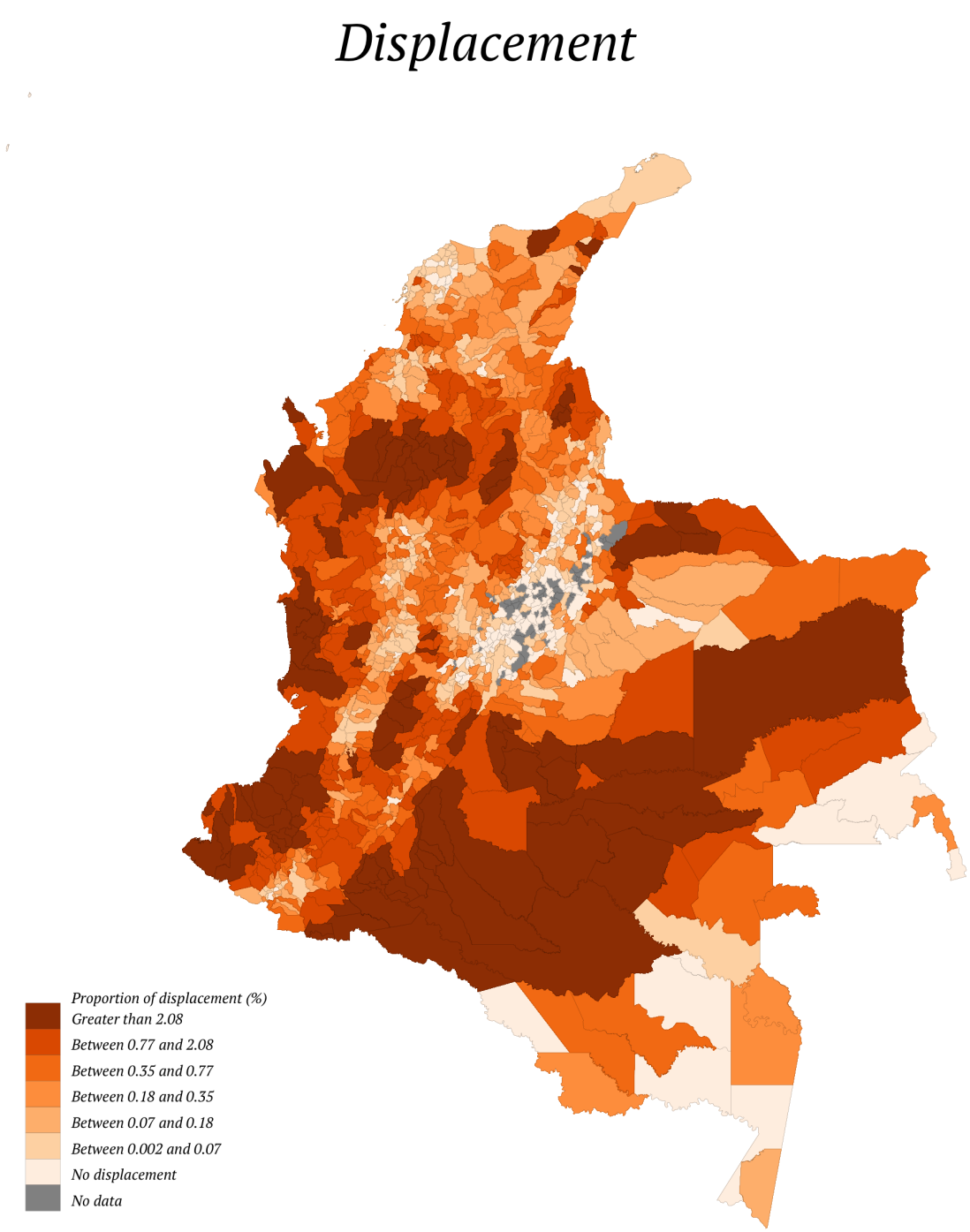
	disp 2010	road index 2010	Pop density	More victims
disp 2010	1.0000			
road index 2010	-0.1808*	1.0000		
Pop density	-0.0960*	0.0718*	1.0000	
More victims	0.6496*	-0.0157	-0.0722*	1.0000

Notes: * Statistically Significant at 5% level

[20]The highest municipality is at 3087 meters above sea level (see table 1

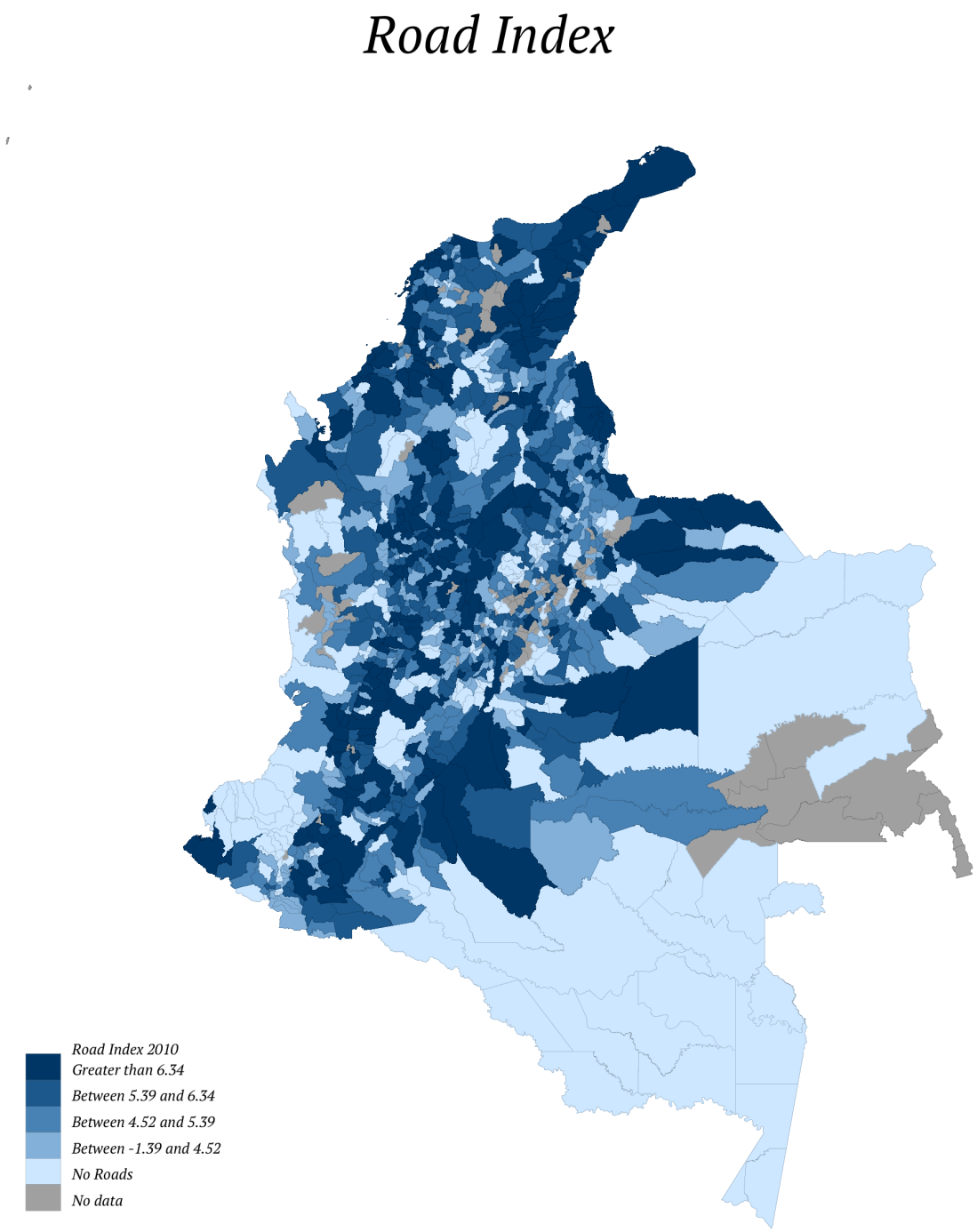
[21]Bogotá, Cali, Medellín, Barranquilla and Cucuta

Figure 2: Map of Displacement rate 2010



Source: Registro unico de victimas

Figure 3: Map of the Road Index 2010



Source: Duranton (2015)

5 Empirical model

The objective of this study is to estimate the impact of roads endowment on conflict, in particular on displacement. For this purpose, let us consider the function:

$$Y_i = f(Roads_i, P_i) \cdot g(Z_i) \tag{1}$$

According to Melo et al. (2012) study, the most common functional form in transport infrastructure studies is equation (1). Where Y_i is the variable measuring conflict in the municipality i, $f(Roads_i, P_i)$ is a decreasing function with the inputs $Roads_i$ *and* P_i which represent the roads endowment and population density in municipality i respectively. The term $g(Z_i)$ is a function of various control factors (e.g geographical, resources, regions, conflict). The function $f(\cdot)$ and $g(\cdot)$ could be expressed as:

$$f(Roads_i, P_i) = \beta_t \; lnRoads_i + \beta_p \; lnP_i \;, \;\; g(Z_i) = \sum_k \gamma_k Z_{k,i}$$

Therefore, the econometric specification considered is the following:

$$Y_i = \beta_0 + \beta_t \; lnRoads_i + \beta_p \; lnP_i + \sum_k \gamma_k \; Z_{k,i} + \xi_i \tag{2}$$

The coefficient of interest is β_t which represents the semi-elasticity of conflict with respect to roads endowment. The first step in the strategy is to estimate the model by OLS excluding the population density. The results of the first step enable me to analyze the behavior of the coefficient of interest. Considering these results, I face different problems. First, I cannot control for unobserved local and time heterogeneity since I do not have a panel. Second, the dependent variable is a ratio, so it must be treated carefully. Third, the road endowment regressor is probably endogenous, so the parameter of interest might be biased. The reason for this is that the decision concerning the level of road endowment in a given municipality is likely to be driven by municipalities with better probabilities to develop the economy, so that, municipalities with conflict reduce the incentive of development. A last concerning is whether or not the agglomeration forces capture a part of the road's effect.

To solve the first problem, I analyze the stability of the coefficient including a broad set of geographical, environmental, regional and conflict controls. The inclusion of these allows me to control in somehow for factors that may confound the effect of roads.

Regarding the second concerning, three different estimators are carried out with all the controls. First, the OLS estimator excluding all the municipalities with no displaced in the dependent variable[22], second a fractional logistic specification, and finally a zero-inflated beta model (ZIB).

[22] Kondylis (2010) follows this approach to measure the impact of displacement on labor market outcomes in post-war BiH

According to Papke and Wooldridge (1996) a fractional logit model may be appropriate in the case where the observed variable is continuous between (0,1). A fractional logit model is a generalized linear model with a binomial distribution and a logit link function. It allows dealing with proportions assuming that the zeros occur through the same process as the other proportions, but it only models the mean. The following specification is estimated:

$$E[Y_i|\mathbf{x}] = G(\beta_0 + \beta_t \; lnRoads_i + \beta_p \; lnP_i + \sum_k \gamma_k \; Z_{k,i}) \tag{3}$$

The linear specification assumes $G(\cdot) = y$, while in the non-linear fractional response regression, $G(\cdot)$ is specified as a logistic function, i.e., $G(z) = exp(y)/(1 + exp(y))$.

To account for the excess of zeros in the dependent variable, the ZIB model is estimated. This consists of two parts: a logistic regression model for whether or not the proportion equals 0, it is a process that distinguishes between zeros and non-zero, and a beta model for the positive proportions. On the contrary to the fractional logit model, this assumes a different process governs observations with zeros and positive outcomes.

The main idea of ZIB model is that there is something qualitatively different about regions that do not present displacement than those who do, at least sometimes.

Let Y_i denote the true displacement proportion in municipality i and let y_i denote the observed estimated displacement proportion in municipality i. Proportions may be in the interval $[0, 1]$. Let $p_i^{(0)}$ be the probability that municipality i has an observed rate of 0. Otherwise, the municipality has a probability of $p_i^{(0,1)} = (1 - p_i^{(0)})$ of having a rate drawn from the Beta(a_i, b_i) distribution. In other words, the result of a multinomial trial is used to determine which of the two processes generates an observation.

$$z_i = \begin{cases} 0 & with\ probability\ p_i^{(0)}, \\ \\ \sim Beta(a_i, b_i) & with\ probability\ p_i^{(0,1)} \end{cases}$$

Therefore, a zero-inflated beta regression model is considered with the following specification:

$$logit(\alpha) = \rho_0 + \rho_1 lnRoads_i + \rho_2 \; lnP_i + \sum_k \rho_k \; Z_{k,i} \tag{4}$$

$$log(\phi) = \beta_0 + \beta_1 lnRoads_i + \beta_2 \; lnP_i + \sum_k \beta_k \; Z_{k,i} \tag{5}$$

After analyzing the robustness of the estimate of the coefficients through special treatment of the zeros, a two-stage-least square (2sls) specification is exploited to deal with the reverse causality in the linear specification. The variable road index is instrumentalized with historical instruments as the literature proposes (Baum-Snow, 2007; Percoco, 2016; Duranton et al., 2013; Duranton, 2015). In particular, I use the length of roads in the year 1938 and if the municipality did not have any road crossing by then the distance to a major road. The network of roads in 1938 should not be related to conflict given that by then, the conflict did not exist or had not started yet. However, the network of roads in 2010 should rely on the roads of 1938.

Bearing in mind whether or not the agglomeration forces capture the road's effect, the variable population density is included in the specification. Since it might cause endogeneity problems, historical instruments for population density are included. Therefore, the estimated model is a system of equations:

$$Y_i = \beta_0 + \beta_t \; lnRoads_i + \beta_p \; lnP_i + \sum_k \gamma_k \; Z_{k,i} + \xi_i \tag{6}$$

$$\begin{aligned} ln \; Roads_i = \pi_1 Roads_{i,1938} + \pi_2 Rdistance_{i,1938} + \pi_3 lnP_{i,1951} + \pi_4 lnP_{i,1964} \\ + \pi_5 lnP_{i,1985} + \pi_6 lnP_{i,1993} + \sum_k \delta_k \; Z_{k,i} + \nu_i \end{aligned} \tag{7}$$

$$\begin{aligned} ln \; P_i = \sigma_1 Roads_{i,1938} + \sigma_2 Rdistance_{i,1938} + \sigma_3 lnP_{i,1951} + \sigma_4 lnP_{i,1964} \\ + \sigma_5 lnP_{i,1985} + \sigma_6 lnP_{i,1993} + \sum_k \sigma_k \; Z_{k,i} + \varepsilon_i \end{aligned} \tag{8}$$

Finally, additional robustness check are carried out. First, I use different instruments for the road index and secondly the variable More victims is used as a measurement of conflict. The roads index instruments are the length of km of Caminos Reales and the distance to the nearest Caminos Reales roads from the municipality.

5.1 The instruments

To deal with the problem of endogeneity in the regression, I make use of instruments that measure the level of roads endowment of the municipalities in the past. The rationale for such choice is driven by the fact that current roads and organization of the country might rely on the first plan of roads network of the government or the roads developed in the colonization period.

The first pair of instruments is the length of the 1938 roads in km and the distance to the nearest principal roads, expressed in km. The latter is a negative value if by the municipality was not passing any road and if it was passing it takes the value of zero. The government in power launched the road network plan to incentive internal trade and development since roads were connecting mostly neighboring municipalities located less than 100km (Duranton, 2015). The exclusion restriction is that the 1938 roads were not built to mitigate the conflict. By then, the conflict, armed groups, and cocaine production did not exist. However, it might have determined the current organization and development of the country. After the construction 1938 roads, different projects were extended based on these roads. The road network plan started to define the center and periphery of the territory.

One of the main concerns on instrument validity is whether it meets the relevance condition. Colombia is a relatively young country, in fact, at the beginning of the XX century most of the country was still jungle and the country was connected mainly by trails. During the second half of the XX century, the country rapidly growth from different point of views, and especially new municipalities emerged. So that nowadays may exist a significant number of municipalities that did not exist by then. However, as I already said at least current principal roads might rely on former roads.

The instruments used as robustness check are the length of the Caminos Reales in km and the distance to the nearest major roads in km. Caminos Reales take roots in the era of colonization, These roads were mainly developed to discover the territory. Under this idea, colonial roads were developed under different circumstances than related to the conflict. The relevance condition may be worst in this case, the issue of the weakness of instruments should be paid attention, especially to Caminos Reales' instrument at the moment of discussing the results.

6 Results

6.1 Baseline estimations

Table 3 reports the OLS estimates of the semi-elasticity of displacement rate with respect to transport infrastructure. The first column does not include any control; column 2 includes the regional controls, column 3 adds regional and geographical controls. Column 4 includes the latter plus the natural resources controls, while column 5 is estimated with all the controls. Most of the controls are a set of dummy variables which try to account for the effect of local factors. Column 6 includes all the controls and a dummy variable which represents the zeros in the departments of Boyacá and Cundinamarca, that is, the variable takes the value of 1 if the municipality belongs to Cundinamarca or Boyacá and presents no displacement. One may observe that the coefficient is robust throughout all

the estimations.

Table 3: Baseline estimations (OLS)

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
road index 2010	-0.0006***	-0.0004***	-0.0005***	-0.0006***	-0.0006***	-0.0006***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Dummy CyB						-0.0001
						(0.0007)
Constant	0.0096***	0.0123***	-0.1723***	-0.1638***	-0.1422***	-0.1414***
	(0.0007)	(0.0019)	(0.0329)	(0.0349)	(0.0296)	(0.0308)
Region controls	No	Yes	Yes	Yes	Yes	Yes
Geo controls	No	No	Yes	Yes	Yes	Yes
Resources controls	No	No	No	Yes	Yes	Yes
Conflict controls	No	No	No	No	Yes	Yes
Observations	1,007	1,007	940	940	940	940
R-squared	0.0327	0.1809	0.2803	0.3027	0.444	0.444
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The dependent variable is the number of displaced over the population in the municipality. The variable road index is presented in logs. Regional controls include the five different regions in the country (Caribbean, Pacific, Andean, Orinoquia, and amazonic). Geographical controls include the climate, altitude, soil aptitude, precipitation, erosion, water, area, and location. Resources controls is a set of dummy variable controlling for the presence of natural resources as coal, emeralds, gold, oil or cattle. Conflict controls is a set of dummies controlling for the presence of coca crops and armed groups.

The results in table 3 indicate that the coefficient associated with the road index is negative and statistically significant at 1% level. The coefficient lies in a range between -0.0004 to -0.0006. The inclusion of all controls increases the goodness of fit of the model considerably, from 0.033 up to 0.44. Additionally, the inclusion of the dummy variable for Boyacá and Cundinamarca has not any significant effect on displacement, that is, these municipalities do not follow a different process concerning the rest of the country.

Even though these estimations present promising and consistent results, the problem of endogeneity, the fact that the dependent variable is a proportion and has a significant amount of zeros are not treated yet. Therefore, I present the results treating the problems separately. The strategy is to check whether the linear approach is reasonable or not, if it is, IV in the linear model is the step to follow to correct for endogeneity.

Table 4 reports different estimations for the special treatment of the zeros, all specifications include all the controls. Column 2 presents estimations excluding all the zeros from the sample. The purpose of this estimation is to examine whether or no the zeros follow a different process. This estimation reports a negative coefficient for road index, which is statistically significant at 1% level. The mean of each sample is calculated to compare column 2 with the results of the full sample (column 1), since the mean of the sample without the zeros tends to be larger, bearing this in mind, it is noticeable that

the coefficient is very similar for both estimations.

Column 3 presents the marginal effect of the fractional logit estimation, the coefficient is also negative and statistically significant at 1% level, however, the magnitude of the effect decreases. Columns 4,5 and 6 reports the marginal effects of the zero-inflated beta model. It takes into consideration that different process governs the proportions of zeros. The marginal effects in columns 4 and 6 show a negative and statistically significant relationship between the mean response (proportion of displaced) and the road index. Differently, the marginal effect in column 5 is not statistically significant. The ZIB model is a combination of two different processes, the results or each process are presented separately. Column 4 presents the total marginal effect, column 5 shows the results for the logistic process for whether or not the proportion equals zero and column 6 presents the marginal effects for the proportions between 0 and 1. I interpret these results as roads have an effect on the amount of displaced for municipalities in which there was an effect of the conflict.

Thus, the results in table 4 suggest robustness in the estimator. The effect decreases marginally but remains significant and negative. This enables me to proceed with IV linear model specification, bearing in mind that OLS estimation might overestimate the road's effect marginally.

However, for the case of the variable More victims the coefficient of interest is not statistically significant in most of the specifications, just for the marginal effects of the logit part in ZIB model the coefficient is negative and statistically significant. The reason is that more than the 50% of the observations for the More victims variable are zeros. Hence, the facts that are used to make up the variable More victims are not as representative as displacement. Most of those facts are close to null for the year 2010 except for homicides. Table 9 shows that most of the coefficients are equal to zero and the R-squared is lower than the estimations with displacement rate.

Table 4: Treatment of the zeros

				Zero inflated Beta model -Marginal effects		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	OLS	OLS-No zeros	FL-Marginal effects	Total Marginal[a]	Logit Marginal[b]	Beta marginal[c]
road_index_2010	-0.0006***	-0.0007***	-0.00044***	-0.00030***	-0.00285	-.000332***
	(0.0001)	(0.0001)	(0.00007)	(0.00007)	(0.00311)	(.000069)
Constant	-0.1422***	-0.1397***				
	(0.0296)	(0.0335)				
Displacement mean	.00721	.008446				
$\beta_1/\bar{y}$	-0.0831	-.0828				
Region controls	Yes	Yes	Yes	Yes		
Geo controls	Yes	Yes	Yes	Yes		
Resources controls	Yes	Yes	Yes	Yes		
Conflict controls	Yes	Yes	Yes	Yes		
Observations	940	809	940	940	940	940
R-squared	0.4440	0.4245				

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The dependent variable is the number of displaced over the population in the municipality. The variable road index is presented in logs. Regional controls include the five different regions in the country (Caribbean, Pacific, Andean, Orinoquia, and Amazonic). Geographical controls include the climate, altitude, aptitude, precipitation, erosion, water, area and location. Resources controls is a set of dummy variable controlling for the presence of natural resources as coal, emeralds, gold, oil or cattle. Conflict controls is a set of dummies controlling for the presence of coca crops and armed groups. Column (2) does not include observations with zeros in the displacement ratio. (a) It presents the total marginal effect of the proportion. (b) It presents the marginal effect of the probability of having a value of 0, logistic regression. (c) It presents the marginal effects of the beta model for the proportions conditional on not having a value of 0.

6.2 Endogenity

Regarding the endogeneity problem in the specification, a 2SLS regression is carried out. Following Duranton (2015), I use as instruments the kilometers of length roads in 1938 at the municipality i and the distance in kilometers from municipalities to a major road. The latter allows to include all municipalities that did not have roads or did not exist in 1938. Using previous roads endowment is a novel approach by Baum-Snow (2007); Michaels (2008); Duranton et al. (2013) to deal with the endogeneity problem. In their study of the effect of roads of the US, the authors use the 1528–1850 exploration routes,1898 railroads, and a 1947 plan of the interstate highway network for the US. Similarly, Percoco (2016) uses the geography of Roman roads in his study of Italian highways.

Panel A Table 5 offers the results for the first stage estimation following the first OLS estimations, gradually adding the controls. It presents the coefficients of the instruments which are statistically significant at 1% level. Additionally, the length of roads and distance roads instrument are positive and negative correlated, respectively, as it was expected.

In the bottom of the table, the R-squared and first stage statistics are presented. Even though the R-squared is low, the F-statistic and the weak identification test reject the null hypothesis that the instruments are weak. Panel B of Table 5 reports the results for the second stage estimation and over-identification test. One may notice that the road index coefficient is negative and statistically significant for all the specifications. Additionally, the inclusion of the instruments increases the magnitude of the coefficient substantially. The results of the over-identification test suggest that the instruments are valid, wherein all the cases the null hypothesis is not rejected.

Table 5: Two stage least square estimation

Two Stage Least square					
Panel A Variables	(1)	(2)	(3)	(4)	(5)
length_roads_1938	0.0796***	0.0817***	0.0870***	0.0841***	0.0644***
	(0.0072)	(0.0078)	(0.0092)	(0.0096)	(0.0098)
dist_roads_1938	-0.0140***	-0.0155***	-0.0101***	-0.0103***	-0.0102***
	(0.0014)	(0.0016)	(0.0024)	(0.0025)	(0.0024)
Constant	3.9925***	4.8311***	-25.9215***	-24.4121***	-25.4307***
	(0.1287)	(0.6039)	(8.5825)	(9.0439)	(9.0209)
[1em] Observations	1,007	1,007	940	940	940
R-squared	0.1396	0.1683	0.1893	0.2089	0.2346
First stage statistics					
F-statistic	148.18***	125.12***	77.91***	71.58***	45.77***
Weak identification test	81.44	75.49	44.28	43.41	29.26
Panel B					
road_index_2010	-0.0022***	-0.0009*	-0.0010***	-0.0012***	-0.0012***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0004)
Constant	0.0154***	0.0287***	-0.1748***	-0.1680***	-0.1570***
	(0.0019)	(0.0036)	(0.0335)	(0.0354)	(0.0330)
Region controls	No	Yes	Yes	Yes	Yes
Geo controls	No	No	Yes	Yes	Yes
Resources controls	No	No	No	Yes	Yes
Conflict controls	No	No	No	No	Yes
Observations	1,007	1,007	940	940	940
R-squared	0.13	0.1627	0.2605	0.2770	0.4189
Overid p-value	0.108	0.43	0.79	0.57	0.43

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: The instrumented variable is the road index 2010. Instruments are the number of kilometres of roads, distance to the nearest road for 1938. The dependent variable is the displacement ratio Stock-Yogo weak ID test critical values: 10% maximal IV size = 19.93 ; 10% maximal IV size = 11.59

6.3 Estimations with population density

The instruments used for the road index remain statistically significant and with the previous signs. Additionally, the population density in 1993 is the most accurate instrument for population density in 2010 since the new system of municipalities was defined in 1991, and the exponential development of the cities started in the early 80's. Results are robust

to the exclusion of the population in 1993. However, the value for the weak identification test decreases marginally. The results also suggest a positive correlation between population density and previous road endowment as previous studies have predicted. Column 4 shows that the population density is statistically significant and has a negative effect on displacement, this confirms the hypothesis proposed. Despite these results, the causal effect of roads remains stable, that is, the population density does not capture the road's effect. The roads index coefficient is marginally lower when population density is included. The effect of population density is higher than roads' effect, almost by double. An increase of 10% of the population density decreases the amount of displaced in 2.3%

The results of the weak identification test in table 6 reject the null that the instruments are weak. The null of the over-identification test is not rejected, indicating that the instruments are valid. The null of the F-statistic in the first-stage is also rejected with high confidence above 99%.

Table 8 presents the results of the estimations using Caminos Reales as instruments. The coefficient of interest is not statistically significant, and the weak identification test suggests that instruments are weak.

Table 6: Estimation including population density

	2SLS			
		First-stage		Second-stage
	(1)	(2)	(3)	(4)
VARIABLES	OLS	Road_index	Population densitty	
road_index_2010	-0.0005***			-0.0011**
	(0.0001)			(0.0004)
l_pop_density_2010	-0.0023***			-0.0023***
	(0.0003)			(0.0004)
length_roads_1938		0.0630***	0.0053***	
		(0.0103)	(0.0011)	
dist_roads_1938		-0.0087***	0.0005**	
		(0.0023)	(0.0002)	
l_pop_density_1964		0.0143	-0.0016	
		(0.0453)	(0.0044)	
l_pop_density_1951		0.0469	-0.0101***	
		(0.0406)	(0.0038)	
l_pop_density_1985		-0.0491	0.0012	
		(0.0685)	(0.0062)	
l_pop_density_1993		0.4426***	1.0210***	
		(0.1432)	(0.0137)	
Constant	-0.1796***	-28.1253***	-2.3039**	-0.1983***
	(0.0296)	(10.0895)	(0.9336)	(0.0353)
Region controls	Yes	Yes	Yes	Yes
Geo controls	Yes	Yes	Yes	Yes
Resources controls	Yes	Yes	Yes	Yes
Conflict controls	Yes	Yes	Yes	Yes
Observations	940	849	849	849
R-squared	0.4764	0.2630	0.9465	0.4535
F-statistic		16.54***	1120***	
Weak identification test		16.01	35.80	13.35
Overid p-value				0.19

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: I use the formula $\hat{y} = \frac{(y * (N-1) + 0.5)}{N}$ (Smithson and Verkuilen (2006)) to avoid the missing values in the log of population density transformation.The dependent variable is displacement ratio and all controls are included. The instrumented variables are the road index and the log of population density in 2010. Instruments are the number of kilometers of roads, distance to the nearest road for 1938 and past population densities for the years 1951, 1964, 1993, 1985.
Stock-Yogo weak ID test critical values: 5% maximal IV relative bias = 15.72; 10% maximal IV relative bias = 9.48

7 Discussion

Before analyzing the results, a note on effect size is necessary given the abstract interpretation of the road index. The road index includes the length of roads in the municipality, the number of exits and the number of major roads, if there is no any primary road

passing through the municipality the distance to the closest one is negatively taking into account. The critical value of the index is -1.39 which indicates that there are no roads in the municipality connecting with other municipality. Values between -1.39 and 5.39 means that the municipality has a certain level of roads but still poor, and values over 6.34 means well-developed regions in terms of roads, that is with accessibility, a large endowment of roads and with more than one primary road. However, the index does take into account the quality of the roads.

The regression analysis and geographical analysis implies a significant and low level of displacement in regions with better roads endowment. Through all regressions estimated the coefficient for roads varies from -0.0003 (see table 4) up to -0.001 (see table 6), which implies stability regarding the different specifications. Moreover, Population density also has negative effects on conflict, even larger than road's effect. Thus, these results support the necessity of an ambitious project as 4G roads not just for regional development but also to reduce conflict and reinforce the peace agreement.

To quantify the results, one may observe that on average doubling the roads in a municipality may reduce the conflict in 11%. However, for instance, in regions where there are no roads or connectivity the impact is even higher, that is municipalities in departments as Chocó, Nariño, Cauca, Vichada, and Caqueta. Additionally, The development of roads in lagged regions also benefits developed municipalities since the number of displaced received should decrease. Through which mechanisms does this negative effect operate?

The first channel is related with the state capacities named by Walder (1995). The government might have a fast response through coercive capacity, legitimacy and incentive socio-economic development throughout the country. The second is linked with economic outcomes. Roads enhance the internal trade, diversifies the economy, increases productivity, incentives business attraction and participation of the population in the real economy.

Different authors support the latter. Duranton (2015) finds that reducing these pairwise distances by expanding the road system is likely to have significant effects on trade and welfare. He also finds that increasing the stock of roads within-cities causes an increase in the weight and value of exports. Cárdenas and Sandoval (2008) finds that higher road density has a positive effect on productivity, nevertheless, road congestion decreases productivity. Figure 3 shows that the stock of roads is located in the principal cities of the Andean region, in cities in the Caribbean, municipalities in the border with Venezuela and Ecuador. However, municipalities in the halfway cannot participate in the benefits. Therefore, any effort to improve roads (either density or quality) in the main municipalities is going to be offset by the effect of massive displacement since it might increase congestion and disequilibrium in the municipalities. In 2010 the cities of Bogotá, Medellín and Cali received 17.621, 17.273 and 6.787 displaced, respectively.

8 Conclusion

The last 50 years Colombia has suffered the consequences of internal conflict. The massive internal diaspora started in the early 90's, and since then, it has been the most representative fact of the conflict. In 2016 Colombian government signed a peace agreement with the insurgent group FARC, the decision is quite divided in the population, there is still a large part of the population skeptical of it. The signing of the agreement was a huge step. However, it does not ensure future development and long-term peace in the country. Simultaneously to the agreement, the Colombian government has been developing one of its most ambitious project in transport infrastructure (4 generation roads). This ambitious project belongs to the set of channels to build a lasting peace in Colombia.

This paper analyses the impact of transport infrastructure on the conflict, using Colombian data for the year 2010. The results suggest that roads have a significant and large effect on the reduction of conflict, mainly, I find that on average doubling the stock of roads reduces displacement about 11%. Moreover, consistently with previous studies, the projects in transport infrastructure should focus on improving the connectivity of the regions instead of increasing the stock of roads in the municipalities.

This paper also examines the effects of population density on conflict. The results suggest that populated cities act as shelters of the war since it offers the benefits of the agglomeration forces and a feasible way to start a new life. One of the concerning of this study was that the agglomeration's effect captured the road's effect, although the results show that agglomeration might help to decrease the conflict, it does not capture the road's effect.

In summary, the results provide an important novel insight into the transport infrastructure literature. This paper goes through different econometrics specifications to deal with the main problems and confirm consistency in the results. Having this in mind, the hypothesis proposed is confirmed, for the Colombian case, development in roads contributes to decrease conflict substantially. Better roads endowment discourage conflict in general. Unlike different countries with internal conflict, the Colombian conflict takes a particular form due to its geography, financing means and regional diversity. Development in roads is a powerful weapon to defeat conflict and consolidate a lasting peace. Thus, it must be on the agenda of policy-makers to consolidate Colombia as a developed country. The development of roads also may help to redistribute the population throughout the country.

While providing some answers, this paper also raised some issues to be considered for future research. First, having a panel would improve the performance of local factor controls and account for time effects. Second, this study deals with the specific features of the dependent variable and endogeneity separately, using more sophisticated econometrics techniques should allow to treat them jointly. Finally, a more detailed geographical

analysis may offer a better understanding of the treatment of the regions regarding the diversity of the country.

References

Álvarez, A., Arias, C., and Orea, L. (2006). Econometric testing of spatial productivity spillovers from public capital. *Hacienda Pública Española*, 178(3):9–21.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, 23(2):177–200.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Besley, T. and Persson, T. (2010). State Capacity, Conflict, and Development. *Econometrica*, 78(1):1–34.

Boarnet, M. G. (1998). Spillovers and the locational effects of public infrastructure. *Journal of Regional Science*, 38(3):381–400.

Burger, M., van Oort, F., and Linders, G.-J. (2009). On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis*, 4(2):167–190.

Cárdenas, M. (2002). Economic growth in colombia: a reversal of fortune? *Archivos de Economía*, 178.

Cárdenas, M. and Sandoval, C. E. (2008). Transportation infrastructure and productivity: evidence from Colombia.

Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*, 98:158–173.

Castells-Quintana, D. and Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science*, 52(2):343–366.

Centro Nacional de Memoria Histórica (2013). Una guerra prolongada y degradada. Dimensiones y modalidades de violencia. *Basta Ya*, pages 30–109.

Collier, P. and Hoefflery, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56:563–595.

Crescenzi, R. and Rodríguez-Pose, A. (2012). Infrastructure and regional growth in the european union. *Papers in Regional Science*, 91(3):487–513.

Czaika, M. and Kis-Katos, K. (2009). Civil conflict and displacement: village-level determinants of forced migration in aceh. *Journal of Peace Research*, 46(3):399–418.

Donaldson, D. (2010). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. Technical report.

Duranton, G. (2015). Roads and trade in Colombia. *Economics of Transportation*, 4(1-2):16–36.

Duranton, G. (2016). Agglomeration effects in Colombia. *Journal of Regional Science*, 56(2):210–238.

Duranton, G., Morrow, P. M., and Turner, M. A. (2013). Roads and trade: Evidence from the US. *Review of Economic Studies*, 81(2):681–724.

Eisner, R. (1991). Infrastructure and Regional Economic Performance: Comment. *New England Economic Review*, pages 47–58.

Evans, P. and Karras, G. (1994). Are Government Activities Productive? Evidence from a Panel of U.S. States. *The Review of Economics and Statistics*, 76(1):1–11.

Evans, P. B., Rueschemeyer, D., and Skocpol, T. (1985). *Bringing the state back in.* Cambridge University Press.

Fajardo Montaña, D. (2015). Estudio sobre los orígenes del conflicto social armado, razones de su persistencia y sus efectos más profundos en la sociedad Colombiana. In *Contribución al entendimiento del conflicto aramdo en Colombia.*

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

Gramlich, E. M. (1994). Infrastructure Investment: A review essay.

Guevara Rosero, G. C. (2017). The effect of trade on agglomeration within regions. *Romanian Journal of Economic Forecasting –*, XX(1):75–97.

Hitchner, R. B. (2012). Roads, Integration, Connectivity, and Economic Performance in the Roman Empire. In *Highways, Byways, and Road Systems in the Pre-Modern World*, pages 222–234.

Holtz-Eakin, D. (1994). Public-Sector capital and the productivity puzzle. *Economics and Statistics*, 76(1):12–21.

Holtz-Eakin, D. and Schwartz, A. E. (1995). Spatial productivity spillovers from public infrastructure: Evidence from state highways. *International Tax and Public Finance*, 2(3):459–468.

Hulten, C. R. and Schwab, R. M. (1991). Public capital formation and the growth of regional manufacturing industries. *The National Tax Journal*, 44(4):121–134.

Ibáñez, A. M. and Vélez, C. E. (2008). Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia. *World Development*, 36(4):659–676.

Kalmanovitz, S. (2003). *Economía y nación: una breve historia de Colombia.* Editorial Norma.

Kondylis, F. (2010). Conflict displacement and labor market outcomes in post-war Bosnia and Herzegovina. *Journal of Development Economics*, 93(2):235–248.

Masahisa Fujita, B. (2003). Does geographical Agglomaration foster economic growth? And Who gains and loses drom it? *The Japanese Economic Review*, 54(2).

Melo, P. C., Graham, D. J., and Brage-ardao, R. (2012). The productivity of transport infrastructure investment. A meta-analysis.pdf. *1st European Symposium on Quantitative Methods in Transportation Systems*, 43(September):1–4.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4):683–701.

Nuñez, M. P. (2007). Monografía Político Electoral. Departamento de Boyacá. Technical report.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632.

Percoco, M. (2016). Highways, local economic structure and urban development. *Journal of Economic Geography*, 16(5):1035–1054.

Pérez, G. J. (2005). La Insfraestructura Del Transporte Vial Y La Movilización De Carga En Colombia. *Documentos de trabajo sobre Economía Regional*, (64):73.

Presidencia de la Republica de Colombia (2016). Summary of Colombia's Agreement to end Conflict and Build peace. Technical report.

Presidencial, P. and Vicepresidencia, D. I. H. (2007). Diagnóstico Departamental Boyacá Diagnóstico Departamental Boyacá. Technical report.

Saiz, A. (2006). Dictatorships and highways. *Regional Science and Urban Economics*, 36(2):187–206.

Santos Silva, J.M.C, Tenreyro, S. (2006). The log of gravity. *The Review of economics and statistics*, 88(4):641–658.

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11(1):54–71.

Stewart, C. (2013). Zero-inflated beta distribution for modeling the proportions in quantitative fatty acid signature analysis. *Journal of Applied Statistics*, 40(5):985–992.

Unidad para la Atención y Reparación Integral a las Víctimas (2017). Interpretación y Análisis sobre las cifras del Registro Único de Víctimas.

Walder, A. G. (1995). *The waning of the communist state [electronic resource]: economic origins of political decline in China and Hungary.* University of California Press.

Wills, M. E. (2015). Los tres nudos de la guerra colombiana. *Contribución al entendimiento del conflicto armado en Colombia Comisión Histórica del Conflicto y sus Víctimas*, pages 0–44.

Appendix A Definition of the variables

According to Unidad para las victimas (2017) report, these are the definitions of each fact used as a measurement of conflict:

Displacement:A person who has been forced to migrate within the national territory is a victim of forced displacement, abandoning their place of residence or habitual economic activities, because their life, physical integrity, personal security or freedom have been violated or are directly threatened with occasion of serious and manifest violations of international human rights standards, which occurred during the internal armed conflict.

Crimes against sexual integrity It refers to any act or omission aimed at violating the exercise of human, sexual or reproductive rights, aimed at maintaining or requesting sexual, physical or verbal contact or participating in sexual interactions through the use of force or the threat of using it. , intimidation, coercion, blackmail, undue pressure, bribery, manipulation or any other mechanism that nullifies or limits the personal will to decide about sexuality and reproduction.

Forced disappearance: It is the act by which one person subjects another person to deprivation of liberty, whatever the form, followed by his concealment and refusal to acknowledge said deprivation or to give information about his whereabouts, subtracting it from the protection of the law

Homicide: It is a crime that consists of an action or omission by which another person is deprived of life either intentionally or culpably.

Kidnapping: It is the limitation or undue deprivation of the freedom of the person, so that it is impossible to determine in space and time. In space as soon as you can not move, according to your will, from one place to another; and in time, as soon as the action requires permanence to be perfect.

Torture: inflicts physical or psychological pain or suffering on a person, in order to obtain from her or a third party information or confession, to punish her for an act committed by her or that inflicts physical or mental pain or suffering on a person, in order to obtain from her or a third party information or confession, to punish her for an act committed by her or that is suspected of having committed or to intimidate or coerce her for any reason behave some type of discrimination (...) Torture shall not be understood as pain or suffering arising solely from legal sanctions or that are a normal or inherent consequence of them.

Seclusion of children: It refers to norms, jurisprudence and national and international treaties that prohibit the recruitment and linking of children and adolescents to regular and irregular armed groups, and advocate for the protection of minors in contexts of armed conflict.

Table 7: Definition of the variables

VARIABLES	Definition of the variable
Main variables	
disp_2010	Proportion of displaced
Displacement	Number of Displaced
road index 2010	It is a function of the logs of road frequency (Number of principal roads), number of exits from the municipality by principal roads and road lengths(mileage of these roads)
Pop density	$inhabitants/km^2$ in 2010
More victims	It is proportion.It is defined as the sum of the population that has been a victim of crimes against sexual integrity, forced disappearance, homicide, kidnapping, physical personal injuries, tortures and Linking Children and Adolescents to armed groups over the population in the municipality
Instruments	
length roads 1938	Km of roads 1938
dist roads 1938	Distance from the municipality to the major road in Km.
Length Caminos Reales	Km of Caminos reales
Distance Caminos reales	Distance from the municipality to the major road in Km
Pop_dens_1993	$inhabitants/km^2$ in 1993
Pop_dens_1985	$inhabitants/km^2$ in 1985
Pop_dens_1964	$inhabitants/km^2$ in 1964
Pop_dens_1951	$inhabitants/km^2$ in 1952
Regional controls	
Andean	It is a dummy that takes the value of 1 if the municipality belongs to the Andean Region
Caribbean	It is a dummy that takes the value of 1 if the municipality belongs to the Caribbean Region
Pacific	It is a dummy that takes the value of 1 if the municipality belongs to the Pacific Region
Orinoquia	It is a dummy that takes the value of 1 if the municipality belongs to the Orinoquia Region
amazonic	It is a dummy that takes the value of 1 if the municipality belongs to the Amazonian Region
Geographical controls	
Altitude	It is expressed in meters above the sea level
climate	It is a categorical variable; Hot=1, Cold=2, Temperate=3
area	it is expressed in km^2
water	It is the level of water in the municipality (cubic meters)
aptitude	Aptitude of the soil. it follows the 8 FAO categories
precipitation	It is the level of precipitation, measured in mm.
erosion	Erosion of the soil
Natural resources controls	
dummy coal	It is a dummy variable that takes the value of 1 if in the municipality there is extraction of coal
dummy emeralds	It is a dummy variable that takes the value of 1 if in the municipality there is extraction of emeralds
dummy golds	It is a dummy variable that takes the value of 1 if in the municipality there is extraction of gold
dummy oil	It is a dummy variable that takes the value of 1 if in the municipality there is extraction of oil
dummy cattle	It is a dummy variable that takes the value of 1 if in the municipality there is cattle
Conflict controls	
COCA 2010	It is a dummy variable that takes the value of 1 if in the municipality there are coca crops
ELN 2010	It is a dummy variable that takes the value of 1 if in the municipality there is presence of ELN
FARC 2010	It is a dummy variable that takes the value of 1 if in the municipality there is presence of FARC
AUC 2010	It is a dummy variable that takes the value of 1 if in the municipality there is presence of AUC or Bacrim

Appendix B Graphs and maps

Figure 4: Histogram

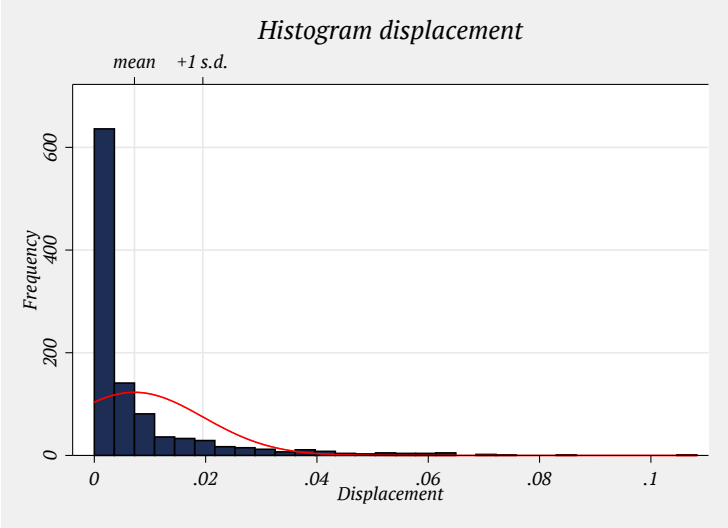


Figure 5: Population Density 2010

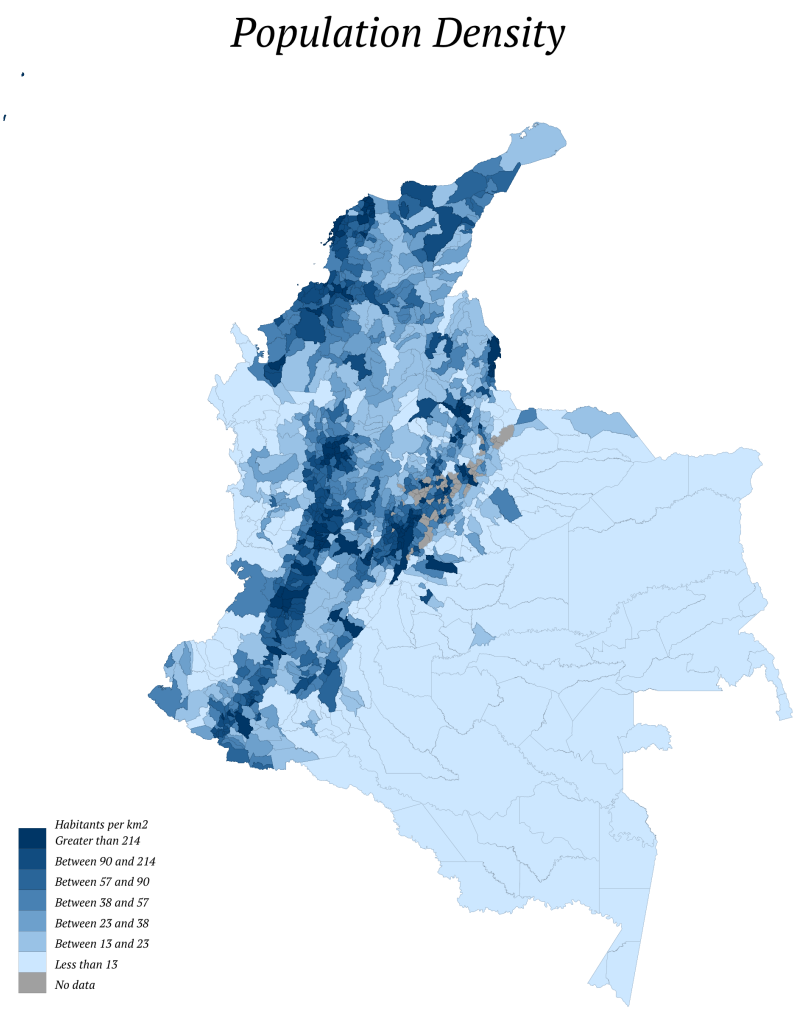
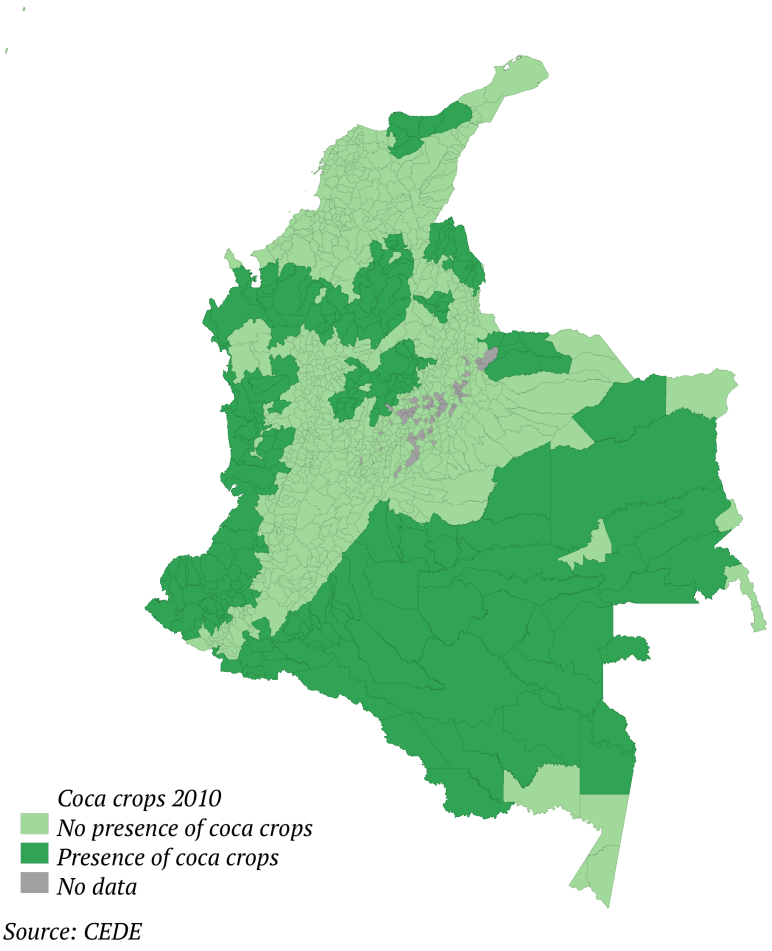


Figure 6: Map of coca crops 2010



Appendix C Tables

Table 8: Estimation including population density with caminos reales

	2SLS		
	First-stage		Second-stage
	(1)	(2)	(3)
VARIABLES	Road_index	Population densitty	Displacement
road_index_2010			-0.0006
			(0.0006)
l_pop_density_2010			-0.0026***
			(0.0004)
length_roads_caminos	0.0336***	0.0018*	
	(0.0086)	(0.0010)	
dist_roads_caminos	-0.0080***	0.0007***	
	(0.0021)	(0.0002)	
l_pop_density_1964	0.0033	-0.0019	
	(0.0456)	(0.0044)	
l_pop_density_1951	0.0503	-0.0098***	
	(0.0404)	(0.0038)	
l_pop_density_1985	-0.0538	0.0015	
	(0.0696)	(0.0062)	
l_pop_density_1993	0.5296***	1.0251***	
	(0.1451)	(0.0135)	
Constant	-25.8102**	-2.3553***	-0.1838***
	(10.3040)	(0.8982)	(0.0372)
Region controls	Yes	Yes	Yes
Geo controls	Yes	Yes	Yes
Resources controls	Yes	Yes	Yes
Conflict controls	Yes	Yes	Yes
Observations	849	849	849
R-squared	0.2471	0.9460	0.4710
F-statistic	10.33***	1123***	
Weak identification test	9.14	20.54	7.59
Overid p-value			0.13

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: I use the formula $\hat{y} = \frac{(y * (N-1) + 0.5)}{N}$ (Smithson and Verkuilen (2006)) to avoid the missing values in the log of population density transformation. Instruments are the number of kilometres of roads, distance to the nearest road for Caminos Reales and past population densities for the years 1951, 1964, 1993, 1985

Table 9: Estimation of Robustness check

	Robustness check						
				Zero inflated Beta model -Marginal effects			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
More victims	OLS	OLS-No zeros	FL-Margins	total-Margins(a)	Logit-Margins(b)	Beta-Margins(c)	Second-stage (2SLS)
road_index_2010	0.0000	-0.0000	-0.0000	0.0000	-0.0162***	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(5.80e-06)	(0.0038)	(0.0000)	(0.00004)
l_pop_density_2010							-.00010***
							(0.00003)
Constant	-0.0137***	-0.0105***					-0.0152***
	(0.0024)	(0.0034)					(0.0029)
Region controls	Yes	Yes	Yes	Yes			Yes
Geo controls	Yes	Yes	Yes	Yes			Yes
Resources controls	Yes	Yes	Yes	Yes			Yes
Conflict controls	Yes	Yes	Yes	Yes			Yes
Observations	940	548	940	940	940	940	849
R-squared	0.3327	0.2392					0.2870
Weak identification test							13.33
Overid p-value							0.65

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: I use the formula $\hat{y} = \frac{(y * (N-1) + 0.5)}{N}$ (Smithson and Verkuilen (2006)) to avoid the missing values in the log of population density transformation. Instruments are the number of kilometres of roads, distance to the nearest road for 1938 roads and past population densities for the years 1951, 1964, 1993, 1985. Column (2) does not include the zeros of the sample. (a) It presents the total marginal effect. (b) It presents the marginal effect of the probability of having a value of 0. (c) It presents the marginal effects of the proportion conditional of not having a value of 0
The variable More victims is a proportion. It is defined as the sum of the population that has been a victim of crimes against sexual integrity, forced disappearance, homicide, kidnapping, physical personal injuries, tortures and seclusion of Children and Adolescents to armed groups over the population in the municipality.