No. **35**
OCTUBRE DE 2015

Documentos **CEDE**

ISSN 1657-7191 Edición electrónica.

Edición especial
CESED

Conflict, Educational Attainment
and Structural Transformation:
La Violencia in Colombia

Leopoldo Fergusson
Ana María Ibáñez
Juan Felipe Riaño

CESED
Centro de Estudios sobre Seguridad y Drogas

CEDE
CENTRO DE ESTUDIOS SOBRE DESARROLLO ECONÓMICO

Universidad de
los Andes
Facultad de Economía

Serie Documentos Cede, 2015-35
ISSN 1657-7191 Edición electrónica.

Octubre de 2015

© 2015, Universidad de los Andes, Facultad de Economía,
CEDE. Calle 19A No. 1 – 37 Este, Bloque W.
Bogotá, D. C., Colombia Teléfonos: 3394949- 3394999,
extensiones 2400, 2049, 3233
infocede@uniandes.edu.co
<http://economia.uniandes.edu.co>

Impreso en Colombia – Printed in Colombia

La serie de Documentos de Trabajo CEDE se circula con propósitos de discusión y divulgación. Los artículos no han sido evaluados por pares ni sujetos a ningún tipo de evaluación formal por parte del equipo de trabajo del CEDE.

El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y sólo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor, sólo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair use), estén previa y expresamente establecidas, no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular, y no atenten contra la normal explotación de la obra.

Conflict, Educational Attainment and Structural Transformation: *La Violencia* in Colombia*

Leopoldo Fergusson[†] Ana María Ibáñez[‡] Juan Felipe Riaño[§]

Abstract

We examine the long-term impact of violence on educational attainment, with evidence from Colombia's *La Violencia*. Individuals exposed to violence during, and especially before, their schooling years experience a significant and economically meaningful decrease in years of schooling. This impact has consequences beyond human capital accumulation: exposed cohorts engage in activities with less human capital content. Violence thus influenced aggregate development – particularly the process of structural transformation, in which some sectors gain prominence as income increases. The effects result not so much from the direct destruction of physical infrastructure, but from affected households' responses to the hardships of conflict.

Keywords: Civil Conflict, Human Capital, Structural Transformation, Development.

JEL: I21, I24, D74, D78, O10, O40

*We acknowledge financial support from the *Centro de Estudios de Seguridad y Drogas, Cesed*. We thank seminar participants at Universidad del Rosario, the 14th Jan Tinbergen European Peace Science Conference, the 2014 Lacea-Lames Meetings, and fellow grantees from Cesed's first round of research grants for their comments.

[†]Universidad de los Andes, Facultad de Economía, lfergusson@uniandes.edu.co

[‡]Universidad de los Andes, Facultad de Economía, aibanez@uniandes.edu.co

[§]University of British Columbia, Department of Economics, jf.riano@alumni.ubc.ca

Conflicto, Escolaridad, y Cambio Estructural: *La Violencia* en Colombia*

Leopoldo Fergusson[†]

Ana María Ibáñez[‡]

Juan Felipe Riaño[§]

Resumen

Examinamos los efectos de largo plazo de *La Violencia* en Colombia sobre el nivel de educación. Los individuos expuestos a la violencia durante, y especialmente antes, de sus años de escolaridad enfrentan una caída estadística y económicamente significativa en sus años de educación. El impacto va más allá de los efectos sobre acumulación de capital humano: las cohortes expuestas se involucran en actividades con menor contenido de capital humano. La violencia por lo tanto afectó el desarrollo agregado – en particular el proceso de cambio estructural, en que unos sectores ganan preeminencia sobre otros a medida que crece el ingreso. Los efectos provienen, no tanto de la destrucción directa de infraestructura física, sino de las reacciones de los hogares a las dificultades del conflicto.

Keywords: Conflicto Civil, Capital Humano, Cambio Estructural, Desarrollo.

JEL: I21, I24, D74, D78, O10, O40

*Agradecemos la financiación del *Centro de Estudios de Seguridad y Drogas, Cesed*. Agradecemos los comentarios de los asistentes a los seminarios de la Universidad del Rosario, la *14th Jan Tinbergen European Peace Science Conference*, la reunión anual de Lacea-Lames en 2014, y de los demás becarios de la primera convocatoria de trabajos de investigación del Cesed.

[†]Universidad de los Andes, Facultad de Economía, lfergusson@uniandes.edu.co

[‡]Universidad de los Andes, Facultad de Economía, aibanez@uniandes.edu.co

[§]University of British Columbia, Department of Economics, jf.riano@alumni.ubc.ca

1 Introduction

Violent civil conflict is prevalent and widespread. Blattman and Miguel (2010) report that over the last 50 years, one-third of all countries have experienced civil war (violent conflicts with at least 1,000 battle-related deaths per year), and nearly one-half have been involved in internal conflict (with at least 25 annual deaths). Understanding the long-term impact of these conflicts and the mechanisms involved is key to developing strategies and public policies to reduce the costs of violence and mitigate their persistence.

Some effects, like the destruction of physical assets, are relatively visible. Moreover, arguably the most important method of recovering these losses is large financial investments. Other “non-physical” costs of violent conflict, however, are less apparent; it is harder to recover from them, as they involve changes in household decisions and opportunities that may reduce future long-run income. Precisely for this reason, these costs may be the most important ones to try to avoid, and the most crucial to focus on in post-conflict environments.¹

Within the set of non-physical effects, lower human capital accumulation stands out as particularly important. Reduced human capital accumulation may be one of the main long-term effects of early-life exposure to conflict. There are various reasons why conflict may influence human capital accumulation. In addition to the direct effect on public goods provision caused by the disorder, chaos and diversion of resources associated with violence, increased uncertainty may reduce households’ incentives to make long-term investments like education. Also, if violence destroys wealth and directly reduces income, households may be forced to withdraw children from school so they can work. Moreover, individuals may abandon school to engage directly in the conflict. Violence also brings distress that directly affects parenting, influencing children’s development and causing stress to pregnant mothers, which may affect fetal development and have long-term implications.

In this paper, we estimate the effects of violence on human capital accumulation (measured as educational attainment) and explore the possible causal mechanisms. Following a relatively recent and growing literature on civil war, we exploit within-country variation in the incidence of violence using data from Colombia’s political violence in the mid-20th century known as *La Violencia*.

The case of *La Violencia* is particularly useful for studying the implications of conflict. First, it was an episode of intense conflict from the mid 1940s to the early 1950s that likely affected households exposed to violence between the traditional Liberal and Conservative parties. Available estimates (see, for instance, Bushnell (1993, cap. 9), Safford and Palacios (2002, cap. 14), Chacón (2004) and Oquist (1980)) indicate that 100,000 to 200,000 people were killed

¹Similarly, conflict may change collective decisions, for instance if violence erodes “social capital.” While this has also been explored in the literature, it will not be the focus of this paper.

(out of a population of around 11.5 million). Second, using information on years of schooling from the 1973 census, we can measure the impact on birth cohorts that were affected by the violence. Third, looking at the implications of the violence as they are revealed in the 1973 census allows us to assess the long-run effect of the bipartisan violence of the 1940s and 1950s, while avoiding potential contamination from the later increase in guerrilla violence in Colombia (which persists today and increased significantly around the mid-1980s).

To understand our empirical approach, consider Figure 1, which presents the main sequence of historical events. Starting in 1948, bipartisan violence increases. While historians disagree about the exact timing of *La Violencia*, our data allow us to identify whether each municipality was affected by the most acute and intense period of bipartisan violence, from 1948 to 1953, and whether they were affected by “Late Violence” (from 1953 to 1959) or the presence of *lit* bandoleros (a phenomenon of violence that persisted after *La Violencia*). In 1958, the traditional Liberal and Conservative parties signed a power-sharing agreement, the *Frente Nacional* (National Front), which effectively tamed bipartisan violence.

The 1973 national census, 25 years after the beginning of *La Violencia*, contains complete micro-level data, including educational attainment. If we consider two individuals born in 1948 (one born in a municipality exposed to violence and the other in a peaceful municipality), they both should have completed their education by the age of 25 in 1973 (as we will show below, this was typically the case at the time). Hence, a first approximation of the effect of violence entails comparing the educational attainment of these two individuals according to the 1973 census.

Of course, the problem with this comparison is that violent and peaceful municipalities may differ on a number of additional dimensions that are correlated with the incidence of violence and also affect educational attainment (e.g., poverty or the quality of local institutions). The obvious solution to this problem is to run a regression in which we control for these municipality characteristics. However, unless we are able to control for all the relevant factors, a regression based on the comparison above will lead to a biased estimate of the impact of violence.

The ideal experiment for avoiding this omitted variable bias is to compare two similar individuals born in the *same* municipality, when it was affected by violence and when it was not. Since each municipality was either affected by violence or not, this is clearly impossible. However, since in 1973 we observe various cohorts, we can approximate the ideal experiment by comparing the educational attainment of an individual born in 1948 in a violent municipality with that of an individual born in the same place well before the increase in bipartisan violence.

This comparison is also potentially problematic, however, as older cohorts may have systematically different levels of education than younger cohorts. Younger cohorts attain better educational outcomes as the country develops, and this will again bias our estimate of the impact of violence.

Our empirical strategy combines these two experiments to find a more convincing estimate of the impact of violence on human capital. Specifically, we compare the difference in the education of individuals who were young enough to be in school (or deciding whether or not to go to school) during *La Violencia* with that of older individuals from the same municipality who had completed their education by then. Next, we compare the size of this gap or ‘trend’ in human capital accumulation in violent (or ‘treated’) vs. peaceful (or ‘control’) municipalities. In this difference-in-differences estimate, only the combination of the two variations in exposure to violence (by place and date of birth) is treated as exogenous. In other words, the strategy uncovers the real effect of *La Violencia* on human capital accumulation if the gap between the younger and older cohorts in peaceful municipalities is a good approximation of what would have been observed had *La Violencia* not occurred.

This “parallel trends” assumption is fundamentally untestable, since the trend of educational attainment in treated municipalities *absent violence* is unknown. However, a major advantage of the approach is that we can verify its plausibility by exploring whether, among cohorts that had finished schooling when violence broke out, the increase in educational attainment from one cohort to another is correlated with the incidence of *La Violencia*. We find no correlation: treated and control municipalities had parallel trends before *La Violencia*. The hypothesis that they would have continued to trend together had it not been for *La Violencia* is therefore more compelling. Moreover, for extra robustness we focus our analysis on a set of more comparable (neighboring) municipalities, and we control directly for potential differential trends in educational attainment in municipalities with different observable characteristics.

Our main result is that individuals exposed to violence, especially the younger cohorts, have a lower educational attainment. For pre-birth/in utero individuals (those 1 to 2 years before birth in 1948), early childhood (0 to 3 years old in 1948), and preschool (4 to 6 years old in 1938), we find a reduction of 0.2 to 0.3 years of education. Effects on the primary school (7 to 12 years old in 1948) and high school (13 to 17 years old in 1948) cohorts are also present, but the absolute decreases in years of schooling are roughly one-third and one-fifth as large, respectively, as the impact on the younger cohorts. Expressed in percentage terms the differences are smaller, because younger cohorts tend to attain more years of schooling on average, given sustained increases in education. However, even taking this into account, the economic significance of the effect on younger cohorts is notably larger. For instance, individuals in the early childhood stage attained close to 4.8 years of schooling on average, so a fall in 0.3 years of schooling represents roughly 6% of the mean attainment. In contrast, the effect on individuals old enough to be in primary school and high school is relatively smaller (0.12 and 0.06 fewer years of education in our baseline estimates): about 3% and 1.5% of average attainment, respectively. We find no differences in attainment for cohorts older than 17 in 1948, reflecting common trends in educational attainment for cohorts too old to have changed their

schooling decisions due to *La Violencia* – which as noted helps validate our approach.

Comparing the size of these effects with those in the existing literature is not straightforward, given the different measures of violence used (e.g., simple indicator variables for exposure versus intensity measures) and the various outcomes measured (e.g., years of schooling, probability of grade completion). Looking at studies that use approaches more comparable to ours suggests a wide range of effects, depending on context and the affected population (boys versus girls, and the younger/in utero versus older kids). Akresh and de Walque (2008) find an average impact of 0.5 fewer years of schooling due to conflict exposure in Rwanda, an 18.3% drop relative to the mean. Rodríguez and Sánchez (2012) report decreases of 0.4 to 1.4 years of schooling in contemporary Colombia, which correspond to falls of 10 to 20% of the mean. The estimates of Chamarbagwala and Morán (2011) are also very sensitive to the group affected, ranging from 6 to 30%. Akbulut-Yuksel (2014) finds falls of 0.4 to 1.2 years of education in a population with an average of 11 years of schooling, which represents roughly a 4 to 10% decrease. Estimates in León (2012) are close to 3%, with 0.3 fewer years of schooling for Peruvians exposed to violence in a sample that had around 9 years of schooling on average. In short, while not comparable to the largest effects found in previous work, the impact we find is still economically meaningful, especially taking into account the coarse measure of violence we have available and the long-run nature of the impact we study.

Another important (and, to the best of our knowledge, novel) aspect of our analysis is that we also show that this impact had long-term consequences that extended beyond human capital accumulation. In particular, we focus on the working population in 1973, and verify the sector in which each individual is employed. We use the average educational levels of employees in the sector (computed out-of-sample, that is, for individuals who were born in municipalities not affected by violence or their neighbors) as a measure of how human-capital intensive it is. Using this measure as our dependent variable, we show that individuals in cohorts exposed to violence are engaged in activities with less human capital.

Again, the effects are concentrated among the relatively younger cohorts, and the size of the effects is economically meaningful. Cohorts in the early childhood stage take the largest hit in terms of the human capital level of the jobs they are able to secure, with 0.31 fewer years of education as a result of violence. This is 5.5% of the human capital content of the average job for individuals in this cohort. Thus, exposure to violence also stops individuals from moving into the more educated (and therefore likely more productive) sectors. This underscores an impact on aggregate development in Colombia and on regional inequality. By slowing the process of structural transformation that occurs as income grows (Herrendorf, Rogerson, & Valentinyi, 2014), violence influenced Colombia’s development process in affected regions.

We find no significant difference in the type of employment among older cohorts. This result suggests that the direct effect of violence on the sector of employment – such as the

destruction of physical infrastructure of more advanced industries or financial capital flight – are less important in slowing down the structural transformation than the human capital accumulation of the workers themselves. Indeed, if these other mechanisms were crucial, there is no a priori reason to expect the effect to be concentrated solely in individuals during their schooling years when violence broke out.

Our results survive a series of different robustness tests. First, we control directly for trends based on observable municipal variables, like historical political alignment, geographic characteristics, or historical institutions and conflicts. Our main results are largely unchanged, which rules out the possibility that the effects of exposure to violence are truly explained by omitted trends that depend on these observables. Second, since *La Violencia* persisted both because some towns experienced a period of “late violence” and because the partisan factions disbanded in some areas into bandit groups, we control for interactions between cohorts and two available proxies for the persistence of violence: the years of late violence and the presence of banditry groups. Our effects estimates are very similar to our baseline findings, suggesting that *La Violencia* mattered directly and not simply due to conflict persistence. Third, we closely examine the potential biases that could emerge from our sample selection. Our main results concern the impact of conflict on the non-migrant population in violent municipalities vs. neighboring peaceful areas. We provide evidence suggesting that these effects are not likely to be heavily influenced by spillovers from overall disorder in the region, or by individuals selectively moving from treated municipalities to peaceful areas. Fourth, we do not find differential marriage or fertility patterns for women exposed to violence during their fertile years, ruling out the possibility that, rather than capturing the direct impact of violence, our estimates conflate the decision of some types of women to change their fertility decisions due to the violence. Finally, exposure to violence cannot predict the absolute size of the treated cohorts or their sex ratio, suggesting that our estimates are not contaminated by differential survival rates among treated cohorts.

To explore what mechanisms underlie this relationship, we study heterogeneous effects by gender and find that our effects are concentrated on boys. This difference suggests that demand-driven mechanisms partly explain our main result, since supply disruption should affect both boys and girls. To further investigate these mechanisms, we study the probability of having *no* schooling at all, rather than the years of schooling, as the outcome. If schools are destroyed and teachers flee, presumably some children simply cannot enter the education system. We find a zero effect on the probability of having no schooling. This finding is in line with our gender-based findings, underscoring that demand factors are the more likely mechanisms driving our overall effect. It is also consistent with findings in the literature, where more detailed data is available.²

²For instance, Swee (2009) notes that in Bosnia and Herzegovina, “direct mechanisms such as the destruction

Our results are important, as they help us understand the implications of conflict beyond the immediate economic effects. The implications on schooling for affected cohorts are of course irreversible, and we suggest that they matter not just because they affect the welfare of affected families, but because they slow down the entire process of income growth and structural transformation. Our results on the likely mechanisms, albeit more suggestive given the limitations of the historical data, are also important from a policy perspective, as they suggest that in order to attenuate the implications of conflict on long-term welfare, policymakers should focus their attention on helping families maintain the conditions that will let them persist in their educational investments, and not just on rebuilding infrastructure.

All the results discussed thus far refer to non-migrants. We find no effects for individuals who migrated from their municipality of birth. This may partly be due to measurement error in exposure to violence for migrants, as our data do not allow us to establish whether individuals moved before or after violence broke out. But since we observe a similar non-effect for even the very young cohorts that could not have moved before being affected, we conclude that, on average, migration helped attenuate the impact of conflict for these households. While this unveils a positive side of migration, it also underscores the inequities that violence brings about: only households fortunate enough to migrate were able to avoid the costs of violence.

A number of papers have also exploited within-country variation (and variation in birth cohorts, in some cases) to study the consequences of violence. The bulk of them has focused on the implications on health (Bundervoet, Verwimp, & Akresh, 2007; Agüero & Deolailkar, 2012). While some examine the effect of violence on education, they typically focus on the short-run effects. Hence, an additional contribution of our research concerns evaluating the long-term effects of violence on human capital accumulation, which has typically not been the focus of the literature. León (2012), who studies the Peruvian civil war, is one notable exception. In the Colombian case, Duque (2014b) also looks at longer-term effects, but she focuses on the increases in general violence (as captured by the homicide rate) starting in the 1980s with the intensification of guerrilla warfare.³ To our knowledge, this is the first paper to examine the effects of *La Violencia* in Colombia on human capital accumulation, using an empirical strategy that aims to overcome the usual identification problems. Similarly, we are not aware of other papers that explore the implications for structural transformation. Finally, we contribute by examining the impact of violence on the first years of life, where the literature is thinner even though there is mounting evidence of the importance of early-life events and conditions (in

of infrastructure and the out-migration of teachers do not seem to matter; instead, youth soldiering may be a key mechanism.” Several of the papers referred to below also emphasize the importance of demand-side mechanisms.

³Duque also presents a useful review and summary table of some of the recent papers examining the impact of war or conflict on human capital, including years of schooling (Akresh & de Walque, 2008; Shemyakina, 2011; Rodríguez & Sánchez, 2012; Chamarbagwala & Morán, 2011; Akbulut-Yuksel, 2014), school enrollment (Swee, 2009) and grade failure Caudillo and Torche (2014).

particular early-life health, see Currie and Vogl (2013)) for long-run results.

The rest of the paper is organized as follows. Section 2 gives an overview of the historical context. Section 3 presents our empirical approach and data sources. Section 4 presents our main results regarding the impact of violence on years of schooling and structural change, as well as the main robustness checks. Finally, section 6 explores the possible causal mechanisms underlying the main results. The final section concludes.

2 Context: *La Violencia* and its impact on education

During the 20th century, Colombia faced three internal conflicts: the *War of the Thousand Days* from 1899 to 1902, *La Violencia* halfway into the century, and the current conflict, which has historical roots but significant impacts since the mid-1980s. Political power struggles between the two traditional parties (Liberal and Conservative) were central to the first two conflicts, and were particularly strong during *La Violencia*. These national-level political struggles fuelled long-standing local disputes for political control and land access, as well as revenge raids and banditry (Oquist, 1980; Sánchez & Meertens, 1983).

Bipartisan violence intensified in 1948 after the assassination of the liberal presidential candidate, Jorge Eliécer Gaitán, leading to the period commonly referred to as *La Violencia*. Violence dropped sharply after a power-sharing agreement brokered peace in 1958. Yet because the agreement only dealt with the political struggles between the two major parties, the other sources of conflict remained unresolved. Many combatants joined local bandit groups, which were often supported by local politicians who lost power during *La Violencia*, thus generating a renewed cycle of violence between 1958 and 1965. In 1964, 100 bandit groups were active (Sánchez & Meertens, 1983).

Aggression against the civilian population during *La Violencia* was intense. Estimates of the extent of deaths during this period diverge. Oquist (1980) calculates that around 194,000 people died between 1948 and 1966, and Guzmán, Fals-Borda, and Umaña (1964) estimate that 200,000 were killed up to 1962. As a percentage of the population of 11 million, the casualties were higher than the sum of all American deaths during the two world wars and the Korean wars combined (Lipman & Havens, 1965).

In addition to deaths and maiming, the violence also caused several costs including illegal land seizure and economic losses. Due to forced migration and illegal land seizure, a little less than 394,000 hectares of land were lost (Oquist, 1980). Economic losses from idle land and asset destruction were also significant. In Tolima, one of the hardest-hit states during this period, the value of estimated economic losses was close to the total national budget of Colombia for 1955 (Guzmán et al., 1964).

The rural population suffered the bulk of this victimization (Sánchez & Meertens, 1983),

which caused forced migration to urban centers. A quantitative survey of migrants from Tolima in 1960 reports that 37.5% of families had a member killed, 59% identified violence as the main cause of their relocation, and 93% did not want to return because violence persisted (Pineda, 1960). Calculating the magnitude of forced displacement is difficult, since the population census of this period did not ask about the reason for migrating. Nonetheless, urban population levels experienced a sharp increase between the 1930s and 1960s: the percentage of the population living in urban areas was 31% in 1938, 39% in 1951, and 52% in 1964 (Bushnell, 1993). Urban migration was a common phenomenon in most Latin American countries during this period, yet violence also played an important role in Colombia. Schultz (1971) shows that net migration rates in Colombia were positively correlated with politically motivated homicides from 1951 to 1974. The national government created a special office to help forced migrants settle in urban areas (Guzmán et al., 1964).

During *La Violencia*, public education was in its early stages. Teachers were poorly trained: 41% and 78% of urban and rural teachers, respectively, had primary education or less (Havens, 1965). The historical literature suggests that *La Violencia* caused several impacts on human capital accumulation. First, violence reduced the provision of education by destroying schools, killing teachers, and causing the total collapse of the state in certain regions (Canal, 1966; Guzmán et al., 1964; Oquist, 1980). Second, children dropped out of school for several reasons. Some children, particularly boys, abandoned school after the violent death of their father to make money. Reports from a school that was heavily affected by *La Violencia* estimated that 30% of the children had lost their father (Canal, 1966). Other children had dropped out of school due to a fear of being victimized (Canal, 1966). Although migration to urban centers forced children to temporarily drop out of school, some children were able to re-enroll later since urban centers had more schools. Still, Pineda (1960) finds that 42% of migrant children were not enrolled in school.

The psychological consequences of victimization also played a key role in the impact of *La Violencia* on education. On the one hand, fear and post-traumatic stress led children and their families to retreat into their private lives. An experiment conducted by Lipman and Havens (1965) found that victims of violence showed clear signs of stress and depression as well as the parents' desire for their children to live in the past. On the other hand, narratives about children abandoning school and joining armed groups were widespread (Canal, 1966; Guzmán et al., 1964; Sánchez & Meertens, 1983). Many adolescent victims of violence joined local bandit groups (*bandoleros*) after 1958 and claimed their desire for revenge (Sánchez & Meertens, 1983). Adolescents also became combatants after active recruitment from armed groups and sometimes from teachers Guzmán et al. (1964). By the end of the 1960s the government started implementing programs to reduce the psychological effects of violence on students and teachers, and to provide different role models than the bandit leaders in rural areas (Sánchez & Meertens,

1983).

In short, the available historical record suggests that violence could have had an important effect on human capital accumulation, both by reducing the demand for education and by curtailing the supply. It also highlights that the impact was likely different for the migrant and non-migrant populations. Below, we provide systematic evidence of the quantitative importance of these alleged effects.

3 Empirical approach

3.1 Data and key variables

Our main dependent variables come from Colombia's 1973 national census. First, we consider years of schooling as of 1973. The census does not record preschool years, meaning that we only observe years of schooling from age six onwards. We assume a maximum of 17 years of education (which corresponds to 12 years of primary plus secondary schooling and 5 years of professional training). As we show below, this is a quite conservative threshold that applies to the vast majority of Colombians in 1973. Hence, to focus on individuals who have completed their education, in our years of schooling regressions we look only at those aged at least $17 + 6 = 23$ years of age in 1973. Since *La Violencia* started in 1948, 25 years earlier, this implies examining individuals who were -2 years old or older in 1948. In most of our specifications we aggregate these cohorts into groups by stages in the education process (like León (2012)), though we also report findings using the full year-on-year variation. Our youngest cohort includes prebirth/in utero individuals in 1948 (-2 to -1 years, that is, 1-2 years before birth), followed by early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29).

Our second dependent variable is intended to evaluate whether cohorts exposed to violence not only attain lower educational levels but also, if they are employed, find jobs in sectors with a lower average human capital content. To verify this hypothesis, we focus on the working population in 1973, and determine in which of 33 sectors each individual is employed. Next, we calculate the average educational levels of each of the sectors. To reduce the likelihood that this average education is influenced by violence directly, we compute it out-of-sample, that is, for municipalities unaffected by violence or their neighbors.⁴ Finally, we assign to each individual the average years of education of the sector in which he or she works. Thus, our dependent variable *mean schooling in job sector* is a measure of the average human capital intensity of various jobs.

Each sector is depicted in Figure 2 with its corresponding average years of education. Some

⁴This is yet another reason to focus our analysis on municipalities affected by violence and their neighbors.

of the sectors with the highest human capital content are: social and related community (a result likely driven by medical doctors), real estate and business services (lawyers), insurance and financial institutions, and crude petroleum and natural gas production (engineers). Other sectors with a relatively high human capital content include several industrial activities, including: chemicals and rubber and plastic products; manufacture of paper and paper products; and manufacture of metal products, machinery, and equipment. The sectors with the lowest levels of formal education include fishing, agriculture and hunting, and forestry and logging.

This contrast between sectors with high and low human capital content shows that if violence reduced employment in sectors with high human capital content, it essentially curtailed the structural transformation that accompanies the process of economic development. Changes in the composition of employment are a key part of the process of economic development, with both demand- and supply-side factors moving economies from predominantly agricultural sectors to services and manufacturing (for a review of this extensive literature, see Acemoglu, 2009, Chapters 20-21).

The distributions of our two main dependent variables are presented in Figure 3, by cohort. The left panel shows the average years of schooling by age in 1973 (for the population old enough to have completed their education). It is clear that younger cohorts have higher educational attainment than older cohorts, which simply reflects the country's progress in education over time. Thus, those aged 23-24 years in 1973 (or -1 to -2 in 1948) have, on average, 5.22 years of schooling, compared to just 1.08 years of schooling for the oldest cohort (the few individuals aged 95 to 99 in 1973, or 70 to 74 in 1948). This comparison possibly understates the extent of progress, as survivors of this age in 1973 are possibly among the richest individuals in society. Focusing on cohorts among which survival rates are less likely to differ dramatically, we see for instance that those who were 45 to 49 in 1973 (20 to 24 in 1948) have, on average, 3.34 years of education: almost two full years less than the youngest generation born about 25 years later. By this calculation, therefore, the country progressed at a rate of about 2 more years of education in 25 years.

The right panel shows that the average schooling in the job sector follows a similar pattern, increasing almost monotonically with younger generations. For instance, the younger cohort is employed in sectors that, on average, have employees with 5.33 years of schooling. The generation about 25 years older, 20 to 24 in 1948, is instead involved in sectors with 3.74 years of schooling.

Our main independent variable, *La Violencia* incidence, is coded based on *Revista Criminalidad*, an annual publication of the National Police Department of Colombia. We use information from the first and second editions of the publication, in which the police recorded which areas were most affected by bipartisan conflict during 1948-1953 (Policía Nacional de Colombia, 1958, 1959). This source has been used in several studies of *La Violencia*, like Chacón, Robinson,

and Torvik (2011) and Chacón (2004). It is also in line with information from more detailed regional studies of this episode of civil conflict (Oquist, 1980; Roldan, 2002; Fernández, 2010).

The presence of *La Violencia* is then coded at the municipality level with a dummy variable that equals 1 if the municipality was affected by bipartisan conflict during 1948-1953 and 0 otherwise. We correct this measure, taking into account separations and new foundations of municipalities from the start of *La Violencia* (April 9, 1948) to July 1, 1958 (when the police first published this information) to 1973, when we observe individuals in the census. Thus, for instance, municipalities in 1973 that were created from a larger municipality coded as violent in 1958 are also coded as affected by violence. Our sample includes individuals born in “traceable municipalities,” defined as places that we can perfectly identify over time from the start of *La Violencia* to the census. We have a very high success rate in identifying individuals’ municipality of birth and the incidence of violence in such municipalities (almost 98% of the roughly 18.4 million individuals in the 1973 census).⁵

We also use two measures of conflict persistence. First, we use information on late violence incidence by Chacón et al. (2011), who code the number of years of partisan conflict at the municipality level after 1948. Second, we coded information on the presence of *bandoleros*. The information is again from *Revista Criminalidad*, describing municipalities affected by these groups of rebels from 1962 to 1965 (Policía Nacional de Colombia, 1965).

We can observe long-term migration by comparing whether the municipality of residence in 1973 coincides with the municipality of birth. Thus, our long-term migration dummy variable equals 1 if residence is different from birthplace and 0 otherwise. In our baseline regressions, we focus on non-migrants, as it is more likely that these individuals, if born in a violent municipality, were exposed to violence. Yet like León (2012), we cannot take individuals’ migration histories into account. Thus we may wrongly assign exposure to violence to individuals who, having been born and currently living in a violent (peaceful) municipality, nonetheless lived in a peaceful (violent) municipality in the interim. However, if anything this should bias the exercise against finding effects, as some individuals assumed to be treated where in fact away during violence. Below, we discuss other likely implications of focusing on the non-migrant

⁵Appendix Table A-1 shows that about 83.6% of the population was born in peaceful municipalities, and 13.5% in violent areas. Since the foreign born may differ in a number of dimensions from individuals born in Colombia, we exclude them from the analysis. Indeed, the table shows that the average years of education for this population is 3.72, longer than the 3.19 years of the rest of the sample. However, including them as born in areas not exposed to *La Violencia* would likely produce similar results, as they constitute a very small share (0.75%) of the total sample. The remaining 2.15% of the population was born either in a municipality with an unidentifiable code (1.5% were from municipalities that could not be matched over time and other areas, possibly reflecting data entry errors by census officials) or in the former *Comisarías*, large and sparsely populated territories of the east and south of Colombia (0.65%). The *Comisarías* were then not formally constituted as departments and are reported in the census as a single group, so the exact place of birth is unknown and it is impossible to assign the violence condition of their birthplace to this group. This small share of people with an unknown municipality of birth had a lower average level of education than the sample as a whole.

population. We also show the results for the migrant population.

Finally, for controls and to check heterogeneous effects in an attempt to uncover the causal mechanisms underlying our results, we also take into account the individual’s gender (from the 1973 census) and a set of municipality-level variables $\mathbf{C}_m$. We include variables that, according to existing research, are likely to affect the incidence of violence in general (and *La Violencia* in particular), and that could also influence educational attainment. Thus, a first set of controls include geographical variables such as altitude, terrain and precipitation. These variables, especially rough terrain (Hegre & Sambanis, 2006; Aguirre, 2011) and precipitation level (Miguel, Satyanath, & Sergenti, 2004), have been highlighted in the literature on violence. These variables are from Colombia’s national geographic institute, IGAC, and Colombia’s Institute of Hydrology, Meteorology and Environmental Studies, IDEAM.

Also included is a set of historical institutional variables that comprise the earliest municipal foundation date from Bernard and Zambrano (1993), and indigenous population presence from 1535 to 1540 from Fernández (2010). We also add some variables capturing historical local conflict and politics: land conflicts from 1901 to 1931 from LeGrand (1986), and Liberals’ vote share in the 1946 presidential election by collecting the official records from the national electoral authority.

All variables are described, and their sources specified, in Appendix Table A-2. Table 1 presents descriptive statistics for the main dependent variables in our sample of municipalities (years of schooling and the mean schooling in job sector) as well as for individual migration status, gender, and the municipal-level variables used in the analysis. Years of schooling is close to 3.9 on average in all municipalities, 4 on average in peaceful municipalities and falls to 3.4 for individuals exposed to violence. The pattern is similar for the average schooling in job sector: it averages 4.5 in all municipalities, 4.65 in peaceful municipalities and 4.21 in violent ones. Also, within violent municipalities there is more migration (50% of individuals born in violent municipalities resided elsewhere in 1973) than among peaceful ones (36%), hinting that migration is a response to violence. The gender composition, 52% female, is similar across all types of municipalities. Finally, in our sample of 449 municipalities, 38% were affected by *La Violencia*. Table 1 also displays the basic descriptive statistics for the remaining municipal controls and variables used in the robustness checks and regressions to understand the mechanisms.

3.2 Empirical strategy

The empirical strategy follows Duflo (2001) and exploits the differential impact of *La Violencia* across municipalities and cohorts in Colombia. The first three columns in Panel A of Table 2 previews, for years of schooling, our main result and strategy in a simplified setting. The

columns consider two types of municipalities, ‘treated’ or ‘violent,’ and ‘control’ or ‘peaceful.’ The rows, in turn, consider two types of cohorts: young and old. Of these, only the young cohort is young enough in 1948 to be either attending school or considering the possibility. The old cohort, in turn, had already completed its education when violence broke out.

In the table, which focuses on non-migrants, young cohorts are individuals aged -2 to 18 in 1948. Thus, if enrolled in an educational institution, the older individuals in the young cohort are expected to be completing high school. As shown above in our descriptive statistics Table 1, with an average education of under 4 years of schooling, education beyond primary school – let alone professional or technical training beyond high school – was very rare in Colombia at the time, as seen in Figure 4, which shows the distribution of our dependent variables by gender. Indeed, individuals with more than 11 years of schooling are two standard deviations above the average schooling achievement. Hence, this is an ample definition of cohorts still at school or considering the possibility. Old individuals, in turn, are those aged 19 to 23 in 1948, while the very old are aged 23 to 29 in 1948.

The difference in years of schooling for young cohorts in violent and peaceful municipalities (which is statistically significant and amounts to 0.74 years, as seen in the first row of Column 3 in Panel A) is not a good approximation of the treatment effect of violence: it assesses the impact of violence, as well as any other differences between violent and peaceful municipalities that may influence educational attainment even in the absence of violence. This latter component, or ‘selection bias,’ may, however, be approximated using older cohorts. Since those cohorts were too old to change their schooling decisions as a result of violence, presumably the schooling gap for old cohorts between peaceful and violent municipalities (0.49 years in the table) merely represents selection bias. By subtracting the difference in old cohorts from the difference in young cohorts, we therefore find our estimate of the impact of violence on schooling attainment. This difference-in-differences estimate, shown in bold, suggests that *La Violencia* reduced educational attainment by 0.25 years.

Reading the difference-in-differences estimates by comparing the columns highlights the fact that the gap for old cohorts between peaceful and violent municipalities must serve as a good approximation of the selection bias for our strategy to be valid. To further understand the parallel trends assumption, it is also useful to compare the rows. The comparison shows that in both types of municipalities educational attainment is higher for younger than for older cohorts. However, this increase in years of schooling is larger in peaceful municipalities. The difference in these differences (that is, the smaller increase in violent municipalities) can be attributed to violence as long as the increase (the trend) in educational attainment in peaceful municipalities is a good approximation of what would have happened to treated municipalities absent *La Violencia*. This is another way of interpreting the parallel trends assumption required to establish the validity of the approach.

Panel B of Table 1 considers a control or placebo experiment devised to verify the validity of this identification assumption. Since old cohorts are not supposed to change their schooling decisions based on levels of violence, increases in educational attainment should not vary by type of municipality for these cohorts. This panel compares the trends in educational attainment for the old and very old groups, neither of which could have changed their schooling decisions as a result of violence. In line with our identification hypothesis, the difference-in-differences estimates in this case are close to zero (-0.015 years of schooling, just 6% as large as the effect in Panel A) and not statistically significant. While cohorts in peaceful municipalities were better educated than those in violent municipalities, schooling in both regions was trending at the same pace before *La Violencia* broke out. This finding makes our identification assumption more plausible: since both types of municipalities were trending together before violence broke out, it is reasonable to assume they would have continued to do so had it not been for *La Violencia*.

Columns 4 to 6 of Table 1 perform the same exercise for our second main dependent variable, mean schooling in job sector. In Panel A, we again find a significant difference-in-differences estimate of the impact of violence on the types of jobs that individuals are able to obtain. According to our calculations in Column 6, individuals born in violent municipalities obtain jobs with less human capital content than those born in peaceful places. The average schooling in the job sector is 0.2 years lower for individuals born in violent municipalities. The gap between manufacturing of metal products, machinery and equipment, and agriculture and hunting, for instance, is close to 4 years. Thus, this effect is close to 5% of this distance.

In Panel B, again, we test the parallel trends assumption for this dependent variable prior to the outbreak of *La Violencia*. Like in the first three columns for years of schooling, the difference-in-differences coefficient is not statistically significant and is precisely estimated, merely a 0.008 gap in average years of schooling by sector of employment.

Notice, however, a key important difference between our two dependent variables. Years of schooling should not have been affected for those old enough to have made their schooling decisions. Instead, violence could have influenced the type of employment sector for all individuals, not just those of school age. For instance, via destruction or flight of physical and financial capital, some more sophisticated sectors that are highly dependent on capital (human and physical) can be more affected by violence. If this is the case, individuals from all cohorts should be employed in sectors with lower human capital content in violent areas. Even older and relatively well-educated people would have a hard time finding jobs in sectors where the average human capital of employees is higher. However, we find negative effects just for younger individuals: precisely the generations whose schooling decisions were affected by violence. This finding suggests that the main constraint for the sophisticated sectors was not that they were directly destroyed by violence. Instead, it supports the idea that the fall in human capital of

young cohorts exposed to violence hindered these sectors.

These initial results illustrate both our main results and the plausible validity of our identification assumption. However, we can exploit more fully the variation in schooling by running the following regression, which generalizes the type of comparison between cohorts and places of birth that Table 2 illustrates:

$$y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{ik}) \gamma_k + \sum_k (\mathbf{C}_m \times d_{ik}) \delta_k + \mathbf{X}_i \beta_i + \epsilon_{imk}. \quad (1)$$

In (1), i indexes individuals, m municipalities (of birth), and k cohorts. Hence, y_{imk} is outcome y for individual i from cohort k born in municipality m . We include municipality and cohort fixed effects, α_m and λ_k , respectively. A dummy variable, V_m , indicates whether municipality m was exposed to bipartisan violence during 1948-1953. Finally, d_{ik} is a dummy variable that equals 1 if individual i belongs to cohort k in 1948, and 0 otherwise. As noted, we consider two dependent variables y_{imk} . The first, educational attainment, is measured as the years of schooling of individual i . The second represents the level of sophistication of individuals' jobs, or more specifically, the average level of human capital (years of schooling) of the sector of the economy in which they are employed.

The coefficients of interest are the γ_k . They constitute the difference-in-differences estimates of the average impact of *La Violencia* on birth cohort k . Thus in the regression, we compare a whole set of relevant cohorts rather than a pair of cohorts, as in Table 2. The γ_k , for cohorts k in their schooling years, are unbiased measures of the impact of violence on human capital accumulation if there are no omitted time-varying and municipality-specific effects correlated with violence incidence. As noted, we can verify the plausibility of this assumption by checking whether *La Violencia* helps explain the change in educational attainment for cohorts that were too old for their schooling decisions to have been influenced by the violence.

A second advantage of our regression analysis is that we can include individual controls $\mathbf{X}_i$ and, particularly importantly, differential trends parametrized as functions of municipal characteristics. The term $\sum_k (\mathbf{C}_m \times d_{ik})$ controls for these cohort trends in outcome y , parametrized as functions of municipality characteristics $\mathbf{C}_m$. Even with parallel trends for older cohorts, this robustness check is key. The main threat to our identification strategy comes from the possibility that, even absent the breakout of violence, municipalities affected by *La Violencia* would have trended away from peaceful ones for other reasons. By including these interaction terms, we allow municipalities with traits that are presumably connected with violence to behave differentially. If, after allowing for such differential trends, we still find an effect of violence, then this possibility is unlikely and rules out alternative interpretations of our findings. Finally, since educational achievement across different cohorts within a given municipality is likely correlated, we cluster standard errors at the municipality level.

4 Main results

4.1 Years of schooling

We start by estimating (1) without any controls other than the individual's gender, and the set of cohorts k by educational stage specified in the previous section: prebirth/in utero individuals in 1948 (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29) – which acts as the excluded, comparison cohort. The results are in Table 3. Column 1 shows that, when considering the full sample of individuals, none of the interactions between violence and cohorts is significant. This suggests that violence had no effect on human capital accumulation.

However, this non-effect is entirely driven by the migrant population. Column 2 restricts the sample to the migrant population, defined as those residing in 1973 in a place different to their place of birth. The interaction terms between violence incidence and the cohort dummies are typically not significant and are quantitatively small. The sole exception we find is for the university cohort, where we find a positive but marginally significant coefficient for the key interaction term. Thus, the incidence of violence in the place of birth seems to have been inconsequential for migrants⁶.

There are several possible explanations for this. First, since we do not know the migration history, some of these individuals might have been totally unaffected by violence. Indeed, they may have left their place of birth before the start of *La Violencia*. By erroneously assigning exposure, the true treatment effect of violence on these individuals is thus attenuated. This explanation, however, only applies to the relatively older cohorts. Clearly the youngest cohort (in utero), and presumably most of the early childhood cohort, were at the very least born in a municipality affected by violence, even if they migrated soon after. Thus, this first reason might not be the full story. The second possible explanation of the lack of an effect is that migrants are able to make up for any potential disadvantage of being born in a violent area. For example, these displaced households may move to richer areas and places with better public goods provision. If this effect is strong enough, then violence could in fact produce a net *positive* spillover by sending people from backward regions to more modern areas. This is not the case, as migrants from violent areas seem to perform the same as migrants from other areas. But the fact that there is no negative effect does point to a possible compensation occurring in receiving municipalities. Note moreover, that this is not the only possibility: migrants from violent areas could also be discriminated against in their new locations, and experience worse outcomes than non-migrants. In this case, something else would explain their relatively better performance in our regressions. For instance, a third reason could be that there is positive selection into

⁶Our findings are consistent with those of León (2012), who finds smaller effects for migrants in Perú and discusses some of the possible explanations that follow.

migration, with people escaping violence being particularly well educated. This may also partly explain the lack of an effect on migrants, but again it is not the only possibility: there could be negative selection, with the most vulnerable households migrating as a result of violence.

Regardless of the underlying reason, this discussion underscores that a variety of factors make it particularly hard to identify the true effect of violence on schooling for migrants. Instead, we hope to have a more precise estimate of the true effect of violence on non-migrants. When we narrow the comparison to non-migrants in Column 3 (that is, individuals born in a given municipality and residing in that same place in 1973), we instead find that all the relatively younger cohorts (from in utero to high school) were negatively affected by violence. Moreover, notice that the larger effects are concentrated among the relatively younger of these groups. The estimated effect is a fall of 0.3 years of education for individuals in the early childhood stage. By contrast, the impact on those in high school is only 0.06 years. This fall in the effect of violence as we move to older cohorts is consistent with how few individuals were reaching higher levels of education at that time and thus changed their schooling decisions as a result of conflict. Individuals in utero, early childhood, preschool and even primary school were more likely to be attending school than their older counterparts (regardless of the political situation), and this decision could have been altered by major shocks like the breakout of violence. Individuals at more advanced stages of schooling may also have more resiliency when faced by shocks: after all the fact that they remained in school at a time when average educational attainment was much lower reflects a strong decision to pursue more education.

Finally, relative to the comparison group of individuals older than 24 years, those at the university stage have no differential schooling achievement by type of municipality. The interaction coefficient for this cohort is small (-0.03) and not statistically significant. Again, this reflects the common trend in educational attainment for cohorts too old to have changed their schooling decisions because of *La Violencia* and helps validate the empirical approach.

The size of these effects is also economically meaningful. For instance, from Figure 3 we know that individuals in the early childhood stage attained close to 4.8 years of schooling on average. Thus the estimated impact of violence on their human capital accumulation (a fall in 0.31 years of schooling) is close to 6% of the mean attainment. The absolute fall in education as a result of violence for preschool age children is about 10 percent lower (0.27), but since individuals at this age attain on average proportionally less education, the magnitude of the effect is comparable to that of individuals in the early childhood stages. In contrast, the effect on individuals old enough to be in primary school and high school is relatively smaller (0.12 and 0.06 less years of education, respectively) even after considering the relatively lower educational attainment of these cohorts. For these cohorts the effects are close to 3% and 1.5% of average attainment, respectively. Finally, consistent with the growing evidence of the importance of pre-birth and in-utero conditions for human capital accumulation and other development outcomes

(see, among others, Mansour and Rees (2012); Duque (2014a, 2014b)) we find sizable effects for these cohorts, with an absolute fall of 0.29 years of education, which corresponds to 5.55% of the mean in the sample (5.22 years of education).

Focusing on the non-migrant sample, we now estimate (1) (again without any controls other than gender) but for all cohorts $k \in \{-2, -1, 0, \dots, 28, 29\}$, where k represents the age of individuals in 1948 (the excluded cohort in the set of dummy variables d_{ik} is that aged 29). The resulting interaction terms, γ_k are plotted in Figure 5. This allows for a quick graphical inspection of the impact of *La Violencia*; it verifies that our results are not an artifact of our choices when grouping cohorts into categories by educational stage. While the estimates are somewhat noisy for such a level of disaggregation of cohorts, the pattern is clear and in line with our preliminary results in Table 2 and Column 3 of Table 3. Younger cohorts, representing those whose schooling decisions may still be affected, have a lower educational attainment when exposed to violence. The size of the effect is between 0.2 and 0.4 years of schooling for individuals aged -2 to 6. The size of the coefficient falls thereafter for older individuals. It stabilizes near zero and is typically not significant for individuals who were 12 years or older when violence took place. In short, *La Violencia* reduced schooling by up to 0.4 years for individuals in violent municipalities, and this effect does not appear to be explained by pre-existing differential trends between the different types of municipalities.

To summarize, while there are no effects of *La Violencia* when examining the full sample of individuals, this result hides a statistically significant and economically meaningful impact of violence on non-migrants. This difference between migrants and non-migrants is not surprising, as migrants may have means to compensate for the impact of conflict. Moreover, our estimates for the migrant population could be contaminated with more confounders (like positive selection into migration) than those for non-migrants. For these reasons, in what follows we focus on non-migrants, and interpret all our effects as an intention-to-treat impact of violence on this population in Colombia. In section 5.4 we discuss in more detail the possible implications of focusing on non-migrants

4.2 Structural change

The previous section showed evidence that *La Violencia* reduced human capital accumulation. In this section we explore other potential implications of these effects. One way to look at the long-term consequences beyond human capital accumulation is to explore the implications on productivity and the process of development. If individuals in violent municipalities were less educated as a result of exposure to violence, it is likely that the process of development and structural change was slowed down in these areas. Specifically, individuals in these municipalities are ill equipped to move into sectors of the economy that demand better skills. Thus, they

are less likely to move into sectors that demand human capital and have higher levels of labor productivity.

Thus we run our basic regression (1) where the dependent variable is the average years of schooling of the sector in which the individual is employed. In this regression, therefore, a negative γ_k implies that individuals in cohorts k exposed to violence are engaged in activities with lower human capital content than those who were not exposed.

Columns 4 to 6 of Table 3 show the results, controlling only for gender and again considering the entire sample (Column 4), migrants (Column 5) and non-migrants (Column 6). Consistent with our findings for years of education, we find no effect in the entire sample, a result driven by migrants. When looking at non-migrants, it is clear that violence has a significant effect on the human capital content of individual employment for cohorts that, in Column 3, exhibited an effect on education. Thus, exposure to violence also stops individuals from moving into the more educated (and hence presumably more productive) sectors. Of course, this underscores an impact on both aggregate development and regional inequality in Colombia.

As for education, the larger effects are concentrated among the relatively younger cohorts, and the size of the effects is economically meaningful. Cohorts in the early childhood stage take the largest hit in terms of the human capital level of the jobs they are able to secure, with 0.31 fewer years of education as a result of violence. This is 5.5% of the average human capital content of individuals in this cohort, as shown in Figure 3, and roughly one-third of the distance in human capital between the average worker in the agriculture and hunting sectors relative to the manufacturing of food and beverages sectors (Figure 2).

To show that how we group cohorts does not affect the main results, we also estimated (1) using the average schooling in the job sector as the dependent variable (again, without any controls other than gender) but for all cohorts $k \in \{-2, -1, \dots, 28, 29\}$. The resulting interaction terms, γ_k , are plotted in Figure 6 just as in Figure 5 for years of schooling. This simple graphical inspection, albeit with more noisy coefficients at this granular level, verifies that violence decreased the importance of sectors with higher human capital for overall employment. Moreover, the largest effects are in the relatively younger cohorts.

Finally, since we find no significant difference among older cohorts, this effect does not appear to be driven by pre-existing differential trends between the different types of municipalities. Moreover, in principle, the type of employment of older cohorts could have been affected by violence (unlike schooling decisions, which should not be affected by violence for older cohorts). This result suggests that the direct effect of violence on the sector of employment – such as the destruction of physical infrastructure of more advanced industries or financial capital flight – are not as important in slowing down the structural transformation as the human capital accumulation of the workers themselves. Indeed, if these other mechanisms were crucial, there is no *a priori* reason to expect the effect to be concentrated solely in individuals of school age when

violence broke out. This finding is also consistent with other studies showing that countries are able to recover from relatively large shocks to physical capital and infrastructure, like large natural disasters and war (Miguel and Roland (2011) is a well-known study of the Vietnam bombings).

5 Robustness

5.1 Omitted variables

While we have presented supporting evidence for our identification assumption that violent municipalities would have exhibited similar trends as peaceful ones had it not been for *La Violencia*, for extra robustness we now control directly for trends based on observable municipal variables, C_m . This helps us rule out the possibility that the effects of exposure to violence can be explained by omitted trends that depend on these observables. Indeed, it is possible that younger cohorts in municipalities affected by *La Violencia* would have trended apart from those in peaceful ones for reasons other than conflict. By including the interaction terms $\sum_k (\mathbf{C}_m \times d_{ik}) \delta_k$, we allow municipalities with traits C_m – which are presumably connected to violence – to behave differentially. If we still find an effect of violence, this possibility is unlikely, helping us rule out alternative interpretations of our findings.

We start by examining the variables likely to influence violence in general, and *La Violencia* in particular, according to existing research. These are grouped into institutions, geography, historical conflict, and politics (presented in Table 4). As expected, a number of these variables, singled out in the conflict literature in general and in that concerned with *La Violencia* in particular, are significantly different for peaceful and violent municipalities. Peaceful municipalities have more indigenous people, recent foundation dates, less rainfall (and less rough terrain, though not significantly), and exhibit lower vote shares for the Liberals. We control for any time-invariant influence of these features in our regressions by including municipality fixed effects. Indeed, this is one of the strengths of the difference-in-difference approach. However, it is relevant to verify that these characteristics did not have a differential impact across cohorts.

Table 5 reports the results of allowing for differential trends parametrized as functions of these observable characteristics in the years of schooling regression. Column 1 shows our baseline results for comparison, where only gender is controlled for. In Columns 2 to 5, we interact each set of municipal characteristics with the full set of cohort fixed effects. Column 2 interacts the institutional variables, Column 3 the vector of geographical controls, Column 4 the historical conflict variables, and Column 5 the political controls. In Column 6, the entire set of controls from Columns 2 to 5 is interacted with the cohort dummies, thus allowing for a very flexible specification that lets municipalities with a wide range of characteristics potentially

connected to incidence of violence behave differently in terms of educational attainment by cohort.

The results are very similar across columns, not just in terms of statistical significance but also in the magnitude of the coefficients. The overall findings from section 4 thus seem to reflect the impact of violence specifically, and not some other variables correlated with conflict.

In Table 6 we do the same robustness exercise, with mean schooling of the job sector as the dependent variable. The table is structured exactly the same as Table 5. In this table, some of the interactions with violence do seem to be somewhat sensitive to the inclusion of differential trends based on municipal characteristics. In particular, when including the interaction of institutional variables with cohort dummies in Column 2, the coefficient on the interaction between violence and the in utero and high school generations becomes smaller in magnitude and is no longer statistically significant at conventional levels. For individuals from early childhood to primary school, however, the estimation still indicates a negative impact of violence that is both sizable and statistically significant. Moreover, in the remaining columns – and, importantly, in Column 6, where all controls are interacted with the cohort dummies, the results are in line with our baseline results reproduced in Column 1 of the table. The only exception refers to the high school generation, where we find no significant effects once all controls are included.

Thus, for this outcome we must add the caveat that the impact on some generations in our baseline results, specifically the in utero and high school cohorts, could reflect trends for different municipalities rather than violence. However, given the robustness of the results for the early childhood, preschool and primary school groups, we again conclude that the reduction in the level of human capital of these individuals’ jobs likely reflects the impact of violence, and not some other variables correlated with conflict.

5.2 Violence persistence

As discussed in Section 2, *La Violencia* persisted both because in some towns there was a period of “late violence” and because the partisan factions splintered off in some areas into bandit groups. One question that emerges is whether the findings reported above do not so much reflect the period of intense bipartisan violence commonly known as *La Violencia*, but simply the impact of persisting violence thereafter.

In Tables 7 and 8 we show regressions where, in addition to our interaction with the main violence dummy, we also control for interactions between cohorts and two available proxies for the persistence of violence: the years of late violence from Chacón et al. (2011) and the presence of bandit groups in the early 1960s coded from Policía Nacional de Colombia (1965). These controls are clearly “bad controls” (Angrist & Pischke, 2008) in the sense that *La Violencia*

could have caused this subsequent violence, in turn causing a fall in schooling.

Still, it is informative to run these regressions in the following sense: if violence persistence explains our findings, then including these bad controls should bring the size of our estimated effects to zero and eliminate their significance. Instead, if the treatment effects remain largely unchanged, then this suggests that *La Violencia* must have also mattered for reasons other than simply conflict persistence. Indeed, a comparison of Columns 1 and 3 in Table 7 with our baseline results for years of schooling (Column 1 of Table 5) shows that the latter case is exactly what happens: the size and significance are largely unchanged when we control for *bandoleros* in Column 1 or for years of late violence in Column 3. A similar conclusion emerges when we look at the even columns, which include our entire battery of interacted controls, and thus can be compared with the last column of Table 5. Table 8 runs analogous regressions for the average schooling in the job sector as the dependent variable, and can be compared to results in Table 6. The finding is similar: a minor impact on the size and significance of our main treatment interactions.⁷

Thus we conclude that *La Violencia* mattered directly for the upheaval it created during the intense period of bipartisan violence, and not simply because it spurred persistent violence in subsequent years.

5.3 Sample selection

In our main regression tables, we focused on the impact of violence by comparing non-migrants born in violent areas to non-migrants born in neighboring peaceful municipalities. We therefore restricted our sample along two dimensions: first by focusing on non-migrants, and second by focusing on neighboring municipalities. In this section we discuss the implications of these choices.

In section 4.1 we discussed a number of difficulties when studying the impact of violence on migrants, concluding that the true effect on non-migrants is more easy to establish. Nevertheless, migration creates challenges to identify the impacts even on the the non-migrant population. Recall the two main reasons that motivated our decision to focus only on the set of neighboring municipalities to estimate our baseline effects. First, it is more likely that neighboring municipalities are more similar to each other. Thus, while we also control for observable differential trends based on municipality characteristics, this choice helps assuage concerns that effects are being driven by other unobservable characteristics not included in the regression. Second, this choice is important for our structural transformation regressions. Indeed, it enables us to estimate the typical human capital content of each employment sector

⁷To save space in the tables, we omit the coefficients for the interactions of measures of persistence and cohorts – which are typically not significant.

out-of-sample, for a set of municipalities that is less likely to have been affected directly by violence, for instance by violence spillovers from neighboring municipalities.⁸

At the same time, this last observation (and our discussion of migrants) also raise the question of whether our peaceful or ‘control’ municipalities are affected by spillovers from their violent neighbors. Indeed, if violence creates migration, a likely conjecture is that the relatively well-off have fewer constraints and choose to migrate to peaceful neighboring areas. The arrival of these better-qualified individuals could create positive spillovers on those born in these areas, both by increasing the demand for education and by raising the returns on education due to positive externalities from relatively well-educated peers. Moreover, the opposite effects could be observed in the municipality of origin, from which the relatively better educated are leaving. If this is the case, our regressions overestimate the influence of violence on individuals. But we cannot be sure that this is the case: alternatively, the externalities in receiving municipalities may be negative if these individuals instead displace locals from schools, and positive in municipalities of origin if overcrowding in schools falls. To make matters worse, it may also be the case that the relatively less well off are the more likely migrants (if they have fewer instruments with which to protect themselves from violence and displacement), leading, for instance, to negative spillovers in receiving municipalities. Finally, violence in neighboring municipalities may disturb the social order and local economic activity in the overall region, thus decreasing educational attainment in control municipalities and leading to an underestimation of the effects of violence. In short, the impact of this combination of spillover effects is hard to define *a priori*, and may be positive (increasing average educational attainment, thus leading us to overestimate the costs of violence) or negative (with the opposite implication).

In Table 9 we present some regressions that enable us to examine the likelihood of these effects. In Column 1 we verify whether violence in the municipality of birth, or in municipalities neighboring violent places, helps predict migration in the full sample of individuals. As expected, we find that being born in a violent municipality increases the likelihood of migration by almost 15 percentage points. However, the most important result for us is that being born in any neighbor of violent municipalities *does not* increase the likelihood of migration. This suggests that one of the spillovers discussed in the previous paragraph – disturbance caused by *La Violencia* in the overall region and not simply in the municipality singled out as most intensely affected by violence according to our sources – is unlikely to be a major concern. If it was, we should observe individuals migrating out of these areas at least at a higher rate than from other peaceful areas. But the coefficient is not significant, and is small in magnitude at 1.6 percentage points – almost a full order of magnitude smaller than the coefficient on violent

⁸A third reason, less important conceptually but relevant in practice, is that by focusing on a smaller set of municipalities, our sample size is significantly reduced, thus decreasing computation time on an otherwise very large dataset (of around 17 million observations).

municipalities.

In Column 2 we instead focus on migrants only, and explore whether being born in a violent municipality affects the probability of moving to a neighboring location. It does, but negatively. That is, it is relatively unlikely that migrants will move to a neighboring municipality if they were born in a violent place. That the rate of arrival from violent neighbors is particularly low helps diminish the concern that we have spillovers stemming from a relatively large proportion of treated individuals arriving in our control municipalities. The sheer arrival of many people could contaminate the control municipality by creating congestion. But as noted, the problem would be especially complicated if particularly well or poorly educated individuals were thrown out of violent places into our control areas, not just because it may contaminate the control and opens the door for potential externalities, but also because it changes the composition of the treated municipalities. We examine this in Panel B (Column 3), where we verify that migrants have more years of schooling than non-migrants in the overall sample: the coefficient on a migrant dummy in a regression for years of schooling suggests that migrants have almost 1.4 additional years of schooling compared to non-migrants. More importantly, in Column 4 we include the interaction of the migration dummy with the violence dummy for the municipality of birth. The coefficient is not statistically significant, suggesting that migrants from violent areas are not particularly positively (or negatively) selected relative to any regular migrant. Thus, while there are more people leaving violent municipalities than peaceful ones, the ones leaving seem to be just like any other migrant in terms of educational attainment.

The results from Table 9 suggest that our identification of the effects on non-migrants is not likely to be heavily influenced by spillovers from overall disorder in the region, or from individuals selectively moving out of treated municipalities and into treated areas. However, the results are only suggestive and disentangling each of these possible influences is very challenging and not feasible with the available data. But we can further verify the validity of our results by running our baseline specification for years of schooling with the full set of municipalities, instead of restricting the sample to violent municipalities and their neighbors. The results are shown in Figure 7, which depicts the coefficients for our more flexible specification where each cohort dummy (all cohorts $k \in \{-2, -1, 0, \dots, 28, 29\}$, where k is the age of individuals in 1948) is interacted with the violence indicator. It is interesting and reassuring that the estimates are quite similar to those depicted in Figure 5 for our baseline sample of neighbors. The point estimates are somewhat smaller in magnitude with the full sample of municipalities. For instance, for ages -2 to about 4 the effect is in the range of about 0.2 to 0.25 years of education in the full sample, and about 0.25 to 0.35 in the restricted sample of neighbors. However, these differences in the coefficients of both samples are not statistically significant, and the overall pattern of a decreasing impact as cohorts get older – with no effect for individuals aged 12 or more (also validating the parallel trends assumption) holds in both samples.

In sum, the results in this section suggest that we are not severely under- or overestimating the impact of violence by restricting our attention in our main results to the sample of violent municipalities and their neighbors.

5.4 Fertility and survival

In spite of the evidence presented thus far, there may remain a few concerns about our measured impact of violence on schooling.⁹ In this section we discuss the extent to which differential fertility or selective survival may contaminate our results.

First, it may be the case that, rather than capturing the direct impact of violence, our estimates conflate the decision of some types of women to change their fertility decisions in the face of violence. For example, the lower educational attainment of treated cohorts could simply reflect that mostly less-educated mothers are willing to have children in the presence of high levels of violence. We examine this possibility by verifying whether reproduction and marriage patterns differ according to violence exposure for the sample of women we observe in 1973 (focusing on those between the ages of 15 and 45 in 1948, thus of reproductive age when potentially exposed to violence).

The results are presented in Table 10 for an indicator variable for having any children, the number of children (total live births), and an indicator variable that equals 1 for married women as a dependent variable in a regression analogous to our basic regression equation. After controlling for municipality and cohort fixed effects (we group cohorts in 5-year intervals), we find no differential marriage or fertility patterns for women exposed to violence during their fertile years. Interactions of the violence dummy with cohort fixed effects are not statistically significant, and are quantitatively very close to zero. To illustrate, while the likelihood of having any children or of being married is on average larger than 80% (notice the constant in the regressions in Columns 1 and 3), the interaction effects are typically smaller than 1 percentage point, virtually negligible. Similarly, for an average of roughly 5.5 children (Column

⁹Other possible sources of selection bias are unfortunately untestable with our data. For instance, Duque (2014b) tests for the presence of selective sorting by comparing the characteristics of mothers exposed to violence with those who were not exposed. The idea is that if some types of families (or in particular, mothers) exposed to violence are systematically different from those that are not exposed, then the mothers' characteristics should significantly correlate with exposure to violence. Also, Duque is able to examine the type of concerns related to migration discussed in the previous section in greater detail, as in her dataset she can identify the place of birth of each child for each mother. Thus, she can investigate the potential displacement of mothers who face violence shocks during or after their pregnancy. In our dataset, however, these types of exercises are not feasible for two reasons. First, our examination of longer-term results implies that few of our children (in 1948) are (in 1973, as adults) still living with their mothers. Therefore, we have little information (and information only for a selected sample, those choosing to live with their mothers) on mothers' characteristics. Second, for the sample of mothers that we can match, we do not have information about the order and place of birth of their children. This discussion also explains why we cannot control for mother fixed effects, a strategy followed by Duque (2014b) and other studies.

2), the point estimate of the interaction terms ranges from 0.04 to 0.14 in the absolute number of children.

Second, our estimates could also be contaminated by differential survival rates among treated cohorts. For example, our impact on schooling for the prenatal cohort could be underestimated if only sufficiently strong fetuses can survive the prenatal stress associated with violence. In Table 11 we therefore examine whether exposure to violence can predict the absolute size of the treated cohorts, and their sex ratio (as girls are biologically stronger than boys). Panel A tests if the size of each of our baseline cohorts (from beyond university in Column 1 down to in utero in Column 7) is different in treated versus control municipalities, after including our basic controls.¹⁰ The constants in the regression give a rough estimate of average cohort size in peaceful municipalities, and the violence coefficient indicates whether cohorts are larger or smaller in treated municipalities. Notice that there is a large variation in the size of cohorts across municipalities, as the standard error on the average size of cohorts is very large. However, this variation is typically not associated with violence exposure. Only once, for individuals beyond university, is the violence coefficient significant. In this case, the size of the coefficient, a fall in almost 29 individuals, is close to 8% of the constant. The coefficients in the regressions for the other cohorts are, as stated, typically not significant, though we must add the caveat that they are always negative and not very precisely estimated. Panel B, on the other hand, shows more precisely estimated non-results. The boy ratio of roughly 45% in each cohort is not smaller in violent municipalities, and the size of coefficients is close to zero. Indeed, even the sole significant coefficient (for the preschool cohort) is nonetheless quantitatively small: just 0.98 percentage points.

Overall, the evidence that treated cohorts are simply smaller, that there was selective survival by gender, or that they reflect differing mother characteristics, is much weaker than the overall results suggesting that these cohorts were directly affected by violence in their schooling choices and outcomes.

6 Mechanisms: supply or demand?

Our previous results have established the impact of violence on cohorts in their schooling years, uncovering a significant and quantitatively important reduction in human capital accumulation and restriction in the types of jobs for cohorts exposed to violence. But what mechanisms underlie this relationship? In this section, we study the heterogenous effects based on gender, as this may provide some insight into the relative role of demand- and supply-side effects.

Table 12 presents our basic regressions for years of schooling, restricting the sample to boys

¹⁰Since some cohorts include more years than others, we divide cohort size by the number of years per cohort to facilitate interpretation.

(in Columns 1 and 3) and girls (in Columns 2 and 4). In these regressions, we include the full set of controls from Tables 5 and 6.

When examining the impact of years of schooling, we find that the results are concentrated on boys, rather than girls. Indeed, the coefficients for in utero, early childhood, and preschool cohorts are significant and negative for both genders, but roughly twice as large for boys than for girls. Thus, it seems that violence affected both genders in these age groups, but the largest effects were suffered by boys. Moreover, while violence reduced education in the primary school and high school age groups for boys, we find no significant effects for girls in these age groups.

When looking at the average schooling of the job sector, the difference is even clearer. Only boys have significant, negative interactions between the age groups and violence. Interactions for girls are typically not significant, with the sole exception of the in utero individuals, where in fact the coefficient is positive. That is, violence reduced the quality of jobs for boys but not for girls.

One reason for this is that since girls have lower educational attainment, their levels of schooling simply cannot fall as much. This, plus the fact that the labor participation of girls was smaller than that of boys may help explain why violence did not affect the jobs girls obtain.

But there are other, more important, reasons why *La Violencia* may have affected boys more than girls, and these may reveal the mechanisms whereby conflict influenced education and development. On the one hand, boys are more likely to have participated directly in violent events. On the other hand, even when not participating directly, if their fathers were killed, engaged in violence, or their job opportunities or income were reduced by the fighting, it is more likely that boys would have had to withdraw from school to support the family. This withdrawal would have pushed them into the labor market earlier, with the implication of halting their education and obtaining lower-quality jobs.

This difference between boys and girls therefore suggests that demand-driven mechanisms partly explain our main result for years of schooling. Indeed, if all the effects were supply driven (school destruction, teacher migration, and so on), they should affect both boys and girls. This is also in line with the historical literature emphasizing kids dropping out of school to cover for their fathers, or to engage directly in violence, as reviewed in section 2. Hence, these results underline the importance of designing policies to attenuate the impact on affected households in a conflict environment and creating incentives to keep children in school, and not simply maintaining supply.

Another way to explore the mechanisms is to look at the probability of having *no* schooling at all, rather than years of schooling, as the dependent variable. Presumably, if schools are destroyed and this explains a large part of the overall effect on years of schooling, we should observe an effect on the probability of simply having no schooling. We ran our basic regression using an indicator variable that equals 1 for individuals with no schooling as the dependent

variable and found no robust effect for our treated cohorts. Indeed, the only significant coefficients, though not robust to the addition of controls, are negative for the youngest controls. In general, as shown in detail in Appendix Table A-4, we find a zero effect on the probability of having no schooling. This is in line with the findings by sex, since it underscores that demand factors (instead of the direct destruction of supply) are the more likely mechanisms for our overall effect.

7 Conclusions

La Violencia in Colombia had important implications for economic development that extended beyond the most immediate and obvious effects on physical capital. In particular, it reduced the human capital accumulation of cohorts exposed to violence, an effect that persisted 25 years after the break out of bipartisan violence. Moreover, this slowed the growth of human capital-intensive sectors, thus retarding the structural transformation required in the process of development, and arguably exacerbating regional inequalities.

These effects highlight the importance of devising policy strategies to attenuate the impact of violence. Our findings point to the importance of further investigating policies to counteract the decisions that households in situations of distress must make (like increasing child labor participation), and alleviate the harsh conditions that could produce obstacles to sustained educational investment (not just financial, but also constraints on early childhood development). Indeed, our results suggest that demand-driven mechanisms partly explain the reduction in human capital accumulation, so an approach based on restoring the supply of public education, like building schools or increasing the number of teachers, is insufficient.

References

- Acemoglu, D. (2009). *Introduction to modern economic growth*. Princeton: Princeton University Press.
- Agüero, J., & Deolailkar, A. (2012). *Armed conflict and the health of children and adolescents: Evidence from the Rwanda genocide*. (University of California, Riverside)
- Aguirre, A. (2011, December). *The risk of civil conflicts as a determinant of political institutions* (Working Papers Central Bank of Chile No. 649). Central Bank of Chile.
- Akbulut-Yuksel, M. (2014). Children of War: The Long-Run Effects of Large-Scale Physical Destruction and Warfare on Children. *Journal of Human Resources*, 49(3), 634-662.
- Akresh, R., & de Walque, D. (2008, April). Armed conflict and schooling: Evidence from the 1994 Rwandan genocide. *Review of Economics and Statistics*(47).
- Angrist, J., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, New Haven: Princeton University Press.
- Bernard, O., & Zambrano, F. (1993). *Ciudad y territorio: El proceso de poblamiento en Colombia*. Bogotá: Academia de Historia de Bogotá.
- Blattman, C., & Miguel, E. (2010). Civil war. *Journal of Economic Literature*, 48(1), 2-57.
- Bundervoet, T., Verwimp, P., & Akresh, R. (2007, July). Health and civil war in rural Burundi. *Journal of Human Resources*, 44(2)(2951), 536-567.
- Bushnell, D. (1993). *The making of modern Colombia: A nation in spite of itself*. Berkeley, University of California Press.
- Canal, G. (1966). *Estampas y testimonios de La Violencia*. Bogotá, Colombia: Canal Ramírez.
- Caudillo, M. L., & Torche, F. (2014). Exposure to local homicides and early educational achievement in México. *Sociology of Education*, 87(2), 89-105.
- Chacón, M. (2004). Dynamics and determinants of the violence during “La Violencia” in Colombia 1946-1963. *Unpublished Department of Political Science, Yale University*.
- Chacón, M., Robinson, J. A., & Torvik, R. (2011, June). When is democracy an equilibrium? Theory and evidence from Colombia’s La Violencia. *Journal of Conflict Resolution*, 55(3), 366-396.
- Chamarbagwala, R., & Morán, H. E. (2011, January). The human capital consequences of civil war: Evidence from Guatemala. *Journal of Development Economics*, 94(1), 41-61.
- Currie, J., & Vogl, T. (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, 5(1), 1-36.
- Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4), 795-813.
- Duque, V. (2014a). *Early-life conditions, parental investments, and child development: Evidence from a violent conflict* (Unpublished). School of Social Work, Columbia University.
- Duque, V. (2014b). *The hidden costs and lasting legacies of violence on education: Evidence from Colombia* (Unpublished). School of Social Work, Columbia University.
- Fernández, M. (2010). *Violencia y derechos de propiedad: El caso de ‘La Violencia’ en Colombia*. (Master’s Thesis, Universidad de los Andes)
- Guzmán, G., Fals-Borda, O., & Umaña, E. (1964). *La Violencia en Colombia: estudio de un proceso social*. Bogotá: Tercer Mundo.
- Havens, A. (1965). *Education in rural Colombia: An investment in human resources*. Madison: Land Tenure Center, University of Wisconsin.

- Hegre, H., & Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution*, 50(4), 508-535.
- Herrendorf, B., Rogerson, R., & Valentinyi, A. (2014). Chapter 6 - growth and structural transformation. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth* (Vol. 2, p. 855 - 941). Elsevier.
- LeGrand, C. (1986). *Frontier expansion and peasant protest in colombia, 1850-1936*. Albuquerque: University of New Mexico Press.
- León, G. (2012). Civil conflict and human capital accumulation: Long term consequences of political violence in Perú. *Journal of Human Resources*, 47(4), 991-1023.
- Lipman, A., & Havens, A. E. (1965). The Colombian violencia: An ex post facto experiment. *Social Forces*, 44(2), pp. 238-245.
- Mansour, H., & Rees, D. I. (2012). Armed conflict and birth weight: Evidence from the al-aqsa intifada. *Journal of Development Economics*, 99(1), 190 - 199.
- Miguel, E., & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1), 1-15.
- Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4), 725-753.
- Oquist, P. (1980). *Violence, conflict and politics in Colombia*. New York: New York Academic Press.
- Pineda, R. (1960). *El impacto de la violencia en el Tolima: el caso de El Líbano*. Universidad Nacional de Colombia, Departamento de Sociología.
- Policía Nacional de Colombia. (1958). Incidencia de la violencia en colombia. *Revista Criminalidad*, No: 1.
- Policía Nacional de Colombia. (1959). Reporte de actividades. *Revista Criminalidad*, No: 2.
- Policía Nacional de Colombia. (1965). Cuadrillas de bandoleros que operaron en Colombia en los años de 1962 y 1965. *Revista Criminalidad*, No: 7.
- Rodríguez, C., & Sánchez, F. (2012). Armed Conflict Exposure, Human Capital Investments, And Child Labor: Evidence From Colombia. *Defence and Peace Economics*, 23(2), 161-184.
- Roldan, M. (2002). *Blood and fire: La Violencia in Antioquia Colombia, 1946-1953*. Durham, NC: Duke University Press Books.
- Safford, F. P., & Palacios, M. (2002). *Colombia: Fragmented land, divided society*. New York, Oxford University Press.
- Sánchez, G., & Meertens, D. (1983). *Bandoleros, gamonales y campesinos: El caso de la violencia en Colombia*. Bogotá: El Áncora.
- Schultz, T. P. (1971). Rural-urban migration in Colombia. *The Review of Economics and Statistics*, 53(2), pp. 157-163.
- Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, 95, 186 - 200.
- Swee, E. L. (2009, April). *On War and Schooling Attainment: The Case of Bosnia and Herzegovina* (HiCN Working Papers No. 57). Households in Conflict Network.

Figure 1: Sequence of Events
La Violencia in Colombia

**Figure 2: Average Years of Schooling of Employees,
by Sector in 1973**

Source: 1973 Colombian Census.

Note: Average educational levels of employees in each of the sectors listed. The average is computed out-of-sample, that is, for individuals who were born in municipalities not affected by violence or their neighbors.

Figure 3: Years of Schooling and Mean Schooling in Job Sector

Source: 1973 Colombian Census.

Note: Years of schooling, in the left panel, is the average years of schooling (in 1973) of each of the cohorts listed on the horizontal axis (by their age in 1948). Mean Schooling in Job Sector, on the right, is computed only for the working population in 1973 and refers to the average years of education of employees involved in the sector in which the individual works. To reduce the likelihood that this average education is directly influenced by violence, it is computed out-of-sample for individuals who were born in municipalities not affected by violence or their neighbors.

Figure 4: Distribution of Years of Schooling and Mean Schooling in Job Sector

Source: 1973 Colombian Census.

Note: Lines correspond to a univariate kernel density estimation for variables on the horizontal axis. Years of schooling, in the left panel, is observed in 1973 for people older than 6. Mean Schooling in Job Sector, on the right, is computed only for the working population in 1973 and refers to the average years of education of employees involved in the sector in which the individual works. To reduce the likelihood that this average education is directly influenced by violence, it is computed out-of-sample for individuals who were born in municipalities not affected by violence or their neighbors.

**Figure 5: Difference-in-Differences Estimate of the Effect of Violence on Schooling
By Age in 1948**

Note: The figure depicts the difference-in-differences estimates of the impact of violence on years of schooling for each cohort, by age in 1948. It plots coefficients and 90% confidence intervals for γ_k in the following regression model: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{ik}) \gamma_k + \epsilon_{imk}$, where: y_{imk} is years of schooling of individual i from cohort k born in municipality m ; α_m are municipality and λ_k cohort fixed effects; V_m equals 1 if municipality m was exposed to bipartisan violence during 1948-1953; and d_{ik} is a dummy variable that equals 1 if individual i belongs to cohort $k \in \{-2, -1, 0, \dots, 28, 29\}$, where k represents the age of individuals in 1948 (the excluded cohort in the d_{ik} set is that aged 29).

Figure 6: Difference-in-Differences Estimate of the Effect of Violence on Structural Change By Age in 1948

Note: The figure depicts the difference-in-differences estimates of the impact of violence on mean years of schooling of sector of employment for each cohort, by age in 1948. It plots coefficients and 90% confidence intervals for γ_k in the following regression model: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{ik}) \gamma_k + \epsilon_{imk}$, where: y_{imk} is mean schooling in job sector for individual i from cohort k born in municipality m ; α_m are municipality and λ_k cohort fixed effects; V_m equals 1 if municipality m was exposed to bipartisan violence during 1948-1953; and d_{ik} is a dummy variable that equals 1 if individual i belongs to cohort $k \in \{-2, -1, 0, \dots, 28, 29\}$, where k represents the age of individuals in 1948 (the excluded cohort in the set of dummy variables d_{ik} is that aged 29). Mean schooling in job sector is computed only for the working population in 1973 and refers to the average years of education of employees involved in the sector in which the individual works. To reduce the likelihood that this average education is directly influenced by violence, it is computed out-of-sample for individuals who were born in municipalities not affected by violence or their neighbors.

**Figure 7: Difference-in-Differences Estimate of the Effect of Violence on Schooling
By Age in 1948 Using Full Sample of Municipalities**

Note: The figure depicts the difference-in-differences estimates of the impact of violence on years of schooling for each cohort, by age in 1948. It plots coefficients and 90% confidence intervals for γ_k in the following regression model: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{ik}) \gamma_k + \epsilon_{imk}$, where: y_{imk} is years of schooling of individual i from cohort k born in municipality m ; α_m are municipality and λ_k cohort fixed effects; V_m equals 1 if municipality m was exposed to bipartisan violence during 1948-1953; and d_{ik} is a dummy variable that equals 1 if individual i belongs to cohort $k \in \{-2, -1, 0, \dots, 28, 29\}$, where k represents the age of individuals in 1948 (the excluded cohort in the set of dummy variables d_{ik} is that aged 29).

Table 1: Summary Statistics

	Mean	Sd	Min	Max	N
Panel A: Individual-level variables					
Years of Schooling	3.86	3.61	0	18	3,252,965
Mean Schooling in Job Sector	4.52	2.62	0	17	1,758,296
Migration	0.39	0.49	0	1	9,111,826
Sex (Woman = 1)	0.52	0.5	0	1	9,111,156
- Affected by Violence					
Years of Schooling	3.43	3.31	0	18	971,554
Mean Schooling in Job Sector	4.21	2.53	0	16	525,042
Migration	0.49	0.5	0	1	2,491,612
Sex (Woman = 1)	0.52	0.5	0	1	2,491,612
- Not Affected by Violence					
Years of Schooling	4.05	3.72	0	18	2,281,411
Mean Schooling in Job Sector	4.65	2.65	0	17	1,233,254
Migration	0.36	0.48	0	1	6,619,544
Sex (Woman = 1)	0.52	0.5	0	1	6,619,544
Panel B: Municipal-level variables					
Violence	0.379	0.486	0	1	449
- Colonial Institutions					
Indigenous Population	0.414	0.493	0	1	449
Early Foundation Date	1774.94	125.013	1530	1975	449
- Geographical Covariates					
Log (Altitude of Municipality Head)	6.671	1.304	0.693	8.023	447
Rainfall (mm)	1942.881	932.168	640	7750	447
Roughness (Sd Altitude)	490.962	284.04	1.756	1349.764	449
- Historical Conflict					
Conflict of Land Tenure (1901 to 1931)	0.078	0.268	0	1	449
- Politics					
Liberals' Vote Share in 1946	61.371	27.29	0.097	100	374
- Additional Robustness					
Bandoleros	0.196	0.397	0	1	449
Years of Late Violence	0.937	1.56	0	6	381

Note: Descriptive statistics for the main sample, including violent municipalities and their neighbors. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. The average is calculated using the sectors in Figure 2 (International Standard Industrial Classification of All Economic Activities, Rev.2) and is computed out-of-sample, that is, for individuals who were born in municipalities not affected by violence or their neighbors. *Violence* is a dummy variable that equals 1 for municipalities affected by violence during *La Violencia*. *Migration* is a dummy variable that equals 1 if the respondent's residence was different from his municipality of birth. For more details on the construction and sources of all variables, see Appendix Table A-2.

Table 2: Simplified Difference-in-Differences Estimate of the Effect of Violence and Placebo Test for Parallel Trends

Cohort in 1948	Years of Schooling Achieved			Mean Schooling in Job Sector		
	Type of Municipality			Type of Municipality		
	Violent (1)	Peaceful (2)	Difference (3)	Violent (4)	Peaceful (5)	Difference (6)
<i>Panel A: Treatment effect</i>						
Young (-2 to 18)	3.804 (0.004)	4.547 (0.003)	-0.743*** (0.006)	4.532 (0.005)	5.020 (0.003)	-0.488*** (0.006)
Old (19 to 23)	3.006 (0.010)	3.497 (0.007)	-0.492*** (0.013)	3.575 (0.010)	3.868 (0.007)	-0.293*** (0.013)
Difference	-0.798*** (0.012)	-1.050*** (0.009)	-0.251*** (0.015)	-0.957*** (0.013)	-1.152*** (0.009)	-0.195*** (0.016)
<i>Panel B: Placebo test, verifying parallel trends</i>						
Old (19 to 23)	3.006 (0.010)	3.497 (0.007)	-0.492*** (0.013)	3.575 (0.010)	3.868 (0.007)	-0.293*** (0.013)
Very Old (24 to 29)	2.888 (0.010)	3.365 (0.007)	-0.477*** (0.013)	3.341 (0.010)	3.626 (0.007)	-0.285*** (0.013)
Difference	-0.118*** (0.014)	-0.133*** (0.010)	-0.015 (0.019)	-0.234*** (0.014)	-0.242*** (0.010)	-0.008 (0.018)

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. The average is calculated using the sectors in Figure 2 (International Standard Industrial Classification of All Economic Activities, Rev.2) and is computed out-of-sample, that is, for individuals who were born in municipalities not affected by violence or their neighbors. Violent municipalities are those affected by *La Violencia*. Young cohorts are individuals aged -2 to 18 in 1948; old individuals are those aged 19 to 23 in 1948, and the very old were aged 23 to 29 in 1948.

**Table 3: Effects of *La Violencia* on Schooling and Mean Schooling in Job Sector
By Migration Status**

Dependent Variable	Years of Schooling			Mean Schooling in Job Sector		
Violence × Cohort in 1948	(1)	(2)	(3)	(4)	(5)	(6)
Violence × In Utero	-0.0207 (0.0769)	0.0860 (0.0707)	-0.291*** (0.102)	-0.0392 (0.0575)	0.0413 (0.0431)	-0.207** (0.0963)
Violence × Early Childhood	-0.0573 (0.0773)	0.0180 (0.0677)	-0.309*** (0.114)	-0.0824 (0.0583)	-0.0146 (0.0386)	-0.310** (0.123)
Violence × Preschool	-0.0627 (0.0617)	-0.00850 (0.0576)	-0.269*** (0.0958)	-0.0798 (0.0492)	-0.0297 (0.0372)	-0.292*** (0.111)
Violence × Primary School	-0.0505 (0.0377)	-0.0510 (0.0445)	-0.121*** (0.0392)	-0.0547* (0.0320)	-0.0405 (0.0299)	-0.183*** (0.0706)
Violence × High School	-0.0157 (0.0305)	-0.00649 (0.0368)	-0.0647** (0.0295)	-0.0391 (0.0258)	-0.0394 (0.0268)	-0.0830** (0.0397)
Violence × University	0.0275 (0.0219)	0.0477* (0.0280)	-0.0325 (0.0261)	0.0167 (0.0184)	0.00591 (0.0219)	-0.0163 (0.0241)
Sex (Woman = 1)	-0.509*** (0.0585)	-0.676*** (0.0279)	-0.351*** (0.131)	0.702*** (0.0378)	0.264*** (0.0245)	0.955*** (0.152)
Municipality Fixed Effects	✓	✓	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Sample of Individuals	All	Migrants	Non-migrants	All	Migrants	Non-migrants
Observations	2,706,067	1,584,132	1,121,935	1,352,115	797,458	554,657
R-squared	0.184	0.107	0.348	0.194	0.106	0.425

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. The average is calculated using the sectors in Figure 2 (International Standard Industrial Classification of All Economic Activities, Rev.2) and is computed out-of-sample, that is, for individuals who were born in municipalities not affected by violence or their neighbors. Violent municipalities are those affected by *La Violencia*. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort.

Table 4: Differences Between Violent and Non-Violent Municipalities

Variable	Peaceful	Violent	Difference
- Colonial Institutions			
Indigenous Population	0,466	0,329	0,136**
Early Foundation Date	1762,341	1795,618	-33,277**
- Geographical Covariates			
Log (Altitude of Municipality Head)	6,749	6,541	0,209
Rainfall (mm)	1791,658	2191,639	-399,980***
Roughness (Sd Altitude)	480,825	507,597	-26,772
- Historical Conflict			
Conflict of Land Tenure (1901 to 1931)	0,065	0,100	-0,035
- Politics			
Liberals' Vote Share in 1946	56,261	70,850	-14,588***
- Additional Robustness			
Bandoleros	0,151	0,271	-0,120**
Years of Late Violence	0,774	1,241	-0,466**

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Violent municipalities are those affected by *La Violencia*. For more details on the construction and source of all variables, see Appendix A-1.

Table 5: Robustness: Years of Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
Violence × Cohort in 1948	Dep. Variable: Years of Schooling					
Violence × In Utero	-0.291*** (0.102)	-0.182** (0.0909)	-0.234** (0.0978)	-0.297*** (0.104)	-0.353*** (0.106)	-0.264*** (0.0993)
Violence × Early Childhood	-0.309*** (0.114)	-0.179** (0.0757)	-0.249*** (0.0934)	-0.309*** (0.113)	-0.384*** (0.121)	-0.246*** (0.0851)
Violence × Pre School	-0.269*** (0.0958)	-0.180*** (0.0598)	-0.219*** (0.0754)	-0.269*** (0.0948)	-0.321*** (0.107)	-0.200*** (0.0681)
Violence × Primary School	-0.121*** (0.0392)	-0.107*** (0.0364)	-0.103*** (0.0403)	-0.121*** (0.0393)	-0.142*** (0.0435)	-0.113*** (0.0415)
Violence × High School	-0.0647** (0.0295)	-0.0604** (0.0300)	-0.0512* (0.0301)	-0.0650** (0.0293)	-0.0880*** (0.0309)	-0.0765** (0.0318)
Violence × University	-0.0325 (0.0261)	-0.0403 (0.0248)	-0.0327 (0.0270)	-0.0323 (0.0259)	-0.0423 (0.0286)	-0.0442 (0.0289)
Sex (Woman = 1)	-0.351*** (0.131)	-0.351*** (0.131)	-0.352*** (0.132)	-0.351*** (0.131)	-0.359*** (0.137)	-0.360*** (0.137)
Controls × Cohort	No	Institutions	Geography	H.Conflict	Politics	Full
Municipality Fixed Effects	✓	✓	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Observations	1,121,935	1,121,935	1,117,605	1,121,935	1,062,066	1,057,736
R-squared	0.348	0.348	0.349	0.348	0.344	0.345

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. The set of municipal controls interacted with cohort fixed effects are marked in each column and include the following: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals' vote share in the 1946 presidential election (politics). "Full" refers to the entire set of controls from Columns 2 to 5. For more details on the construction and source of all variables, see Appendix A-1.

Table 6: Robustness: Structural Change

	(1)	(2)	(3)	(4)	(5)	(6)
Violence × Cohort in 1948	Dep. Variable: Mean Years of Schooling in Job Sector					
Violence × In Utero	-0.207** (0.0963)	-0.104 (0.0774)	-0.210** (0.0969)	-0.215** (0.0998)	-0.241** (0.101)	-0.156* (0.0869)
Violence × Early Childhood	-0.310** (0.123)	-0.155* (0.0865)	-0.288*** (0.107)	-0.313** (0.125)	-0.363*** (0.129)	-0.215** (0.0944)
Violence × Pre School	-0.292*** (0.111)	-0.154** (0.0721)	-0.258*** (0.0917)	-0.294*** (0.113)	-0.337*** (0.118)	-0.176** (0.0793)
Violence × Primary School	-0.183*** (0.0706)	-0.111** (0.0496)	-0.168*** (0.0631)	-0.185** (0.0728)	-0.212*** (0.0781)	-0.125** (0.0530)
Violence × High School	-0.0830** (0.0397)	-0.0545 (0.0362)	-0.0727* (0.0396)	-0.0841** (0.0409)	-0.0937** (0.0442)	-0.0487 (0.0383)
Violence × University	-0.0163 (0.0241)	-0.0251 (0.0254)	-0.0297 (0.0264)	-0.0171 (0.0240)	-0.0215 (0.0262)	-0.0399 (0.0290)
Sex (Woman = 1)	0.955*** (0.152)	0.952*** (0.152)	0.955*** (0.153)	0.955*** (0.151)	0.944*** (0.154)	0.942*** (0.155)
Controls × Cohort	No	Institutions	Geography	H.Conflict	Politics	Full
Municipality Fixed Effects	✓	✓	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Observations	554,657	554,657	552,589	554,657	526,420	524,352
R-squared	0.425	0.426	0.426	0.425	0.421	0.422

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. The set of municipal controls interacted with cohort fixed effects are marked in each column and include the following: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals' vote share in the 1946 presidential election (politics). "Full" refers to the entire set of controls from Columns 2 to 5. For more details on the construction and source of all variables, see Appendix A-1.

Table 7: Robustness Persistence of Violence: Years of Schooling

	(1)	(2)	(3)	(4)
Violence $\times$ Cohort in 1948	Dep. Variable: Years of Schooling			
Violence $\times$ In Utero	-0.291*** (0.103)	-0.235** (0.0962)	-0.243** (0.0987)	-0.253** (0.0993)
Violence $\times$ Early Childhood	-0.288*** (0.106)	-0.207** (0.0826)	-0.213** (0.0899)	-0.210** (0.0825)
Violence $\times$ Pre School	-0.249*** (0.0879)	-0.169** (0.0655)	-0.167** (0.0657)	-0.152** (0.0623)
Violence $\times$ Primary School	-0.103*** (0.0394)	-0.0872** (0.0402)	-0.0861** (0.0396)	-0.0903** (0.0400)
Violence $\times$ High School	-0.0585* (0.0306)	-0.0653** (0.0321)	-0.0716** (0.0316)	-0.0774** (0.0314)
Violence $\times$ University	-0.0273 (0.0266)	-0.0356 (0.0290)	-0.0403 (0.0256)	-0.0444 (0.0279)
Sex (Woman = 1)	-0.351*** (0.131)	-0.360*** (0.137)	-0.199*** (0.0695)	-0.202*** (0.0715)
Bandoleros $\times$ Cohort	✓	✓	-	-
Years of Late Violence $\times$ Cohort	-	-	✓	✓
Full Controls $\times$ Cohort	-	✓	-	✓
Observations	1,121,935	1,057,736	875,838	848,985
R-squared	0.348	0.345	0.251	0.251

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. “Full” controls refers to the following municipal variables, which are interacted with cohort fixed effects: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals’ vote share in the 1946 presidential election (politics). For more details on the construction and source of all variables, see Appendix A-1.

Table 8: Robustness Persistence of Violence: Structural Change

	(1)	(2)	(3)	(4)
Violence $\times$ Cohort in 1948	Dep. Variable: Mean Schooling in Job Sector			
Violence $\times$ In Utero	-0.214** (0.0971)	-0.137 (0.0867)	-0.137 (0.0885)	-0.119 (0.0836)
Violence $\times$ Early Childhood	-0.301** (0.116)	-0.185** (0.0933)	-0.214** (0.102)	-0.170* (0.0893)
Violence $\times$ Pre School	-0.289*** (0.104)	-0.157** (0.0769)	-0.186** (0.0829)	-0.122* (0.0716)
Violence $\times$ Primary School	-0.189*** (0.0701)	-0.117** (0.0531)	-0.136** (0.0586)	-0.102** (0.0492)
Violence $\times$ High School	-0.0924** (0.0413)	-0.0479 (0.0387)	-0.0698* (0.0397)	-0.0376 (0.0367)
Violence $\times$ University	-0.0232 (0.0256)	-0.0376 (0.0291)	-0.0185 (0.0279)	-0.0270 (0.0285)
Sex (Woman = 1)	0.955*** (0.152)	0.942*** (0.155)	1.134*** (0.142)	1.126*** (0.145)
Bandoleros $\times$ Cohort_k	✓	✓	-	-
Years of Late Violence $\times$ Cohort_k	-	-	✓	✓
Full Controls $\times$ Cohort	-	✓	-	✓
Observations	554,657	524,352	421,430	408,937
R-squared	0.425	0.422	0.358	0.358

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. “Full” controls refers to the following municipal variables, which are interacted with cohort fixed effects: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), liberals’ vote share in the 1946 presidential election (politics). For more details on the construction and source of all variables, see Appendix A-1.

Table 9: Violence, Migration and Education

Panel A: Violence and Migration			Panel B: Migration and Education		
(1)			(3)		
Dep. Variable:	Migrant	Migrant to Neighbor	Dep. Variable:	Years of Schooling Achieved	
<i>Individual was born in...</i>					
Violent Municipality	0.149*** (0.0218)	-0.0427*** (0.0150)	Migrant	1.135*** (0.0591)	1.118*** (0.0669)
Peaceful Municipality Neighbor of Violent	0.0158 (0.0502)		Violence × Migrant		0.121 (0.101)
Constant	0.341*** (0.0181)	0.221*** (0.0105)	Constant	3.069*** (0.0325)	3.066*** (0.0311)
Sample	All Individuals	Migrants	Sample	Older than 6	Older than 6
Municipality Fixed Effects	-	-	Municipality Fixed Effects	✓	✓
Observations	17,862,833	6,567,639	Observations	6,587,411	6,587,411
R-squared	0.011	0.002	R-squared	0.163	0.163

Note: Robust standard errors in parentheses clustered at the municipality level *** p<0.01, ** p<0.05, * p<0.1

Table 10: Selective Fertility in the Baseline Sample

	(1)	(2)	(3)
Dependent variable:	Dummy=1 if had any children	Number of children	Dummy=1 if married
<i>Violence × Age in 1948...</i>			
Violence x 15 to 19	-0.00393 (0.00699)	-0.0547 (0.123)	0.000334 (0.00986)
Violence x 20 to 24	-0.000199 (0.00635)	0.0424 (0.111)	0.0102 (0.00951)
Violence x 25 to 29	-0.000618 (0.00652)	-0.141 (0.105)	-0.00134 (0.00874)
Violence x 30 to 34	0.00362 (0.00668)	0.0473 (0.108)	0.00441 (0.0104)
Violence x 35 to 39	-0.00436 (0.00597)	-0.0900 (0.0957)	-0.00195 (0.00862)
Constant	0.888*** (0.00215)	5.599*** (0.0415)	0.803*** (0.00283)
Municipality fixed effects	✓	✓	✓
Full Controls × Cohort	✓	✓	✓
Observations	265,100	236,854	265,100
R^2	0.016	0.073	0.014

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cohorts defined by age in 1948. “Full” controls refers to the following municipal variables, which are interacted with cohort fixed effects: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals’ vote share in the 1946 presidential elections (politics). For more details on the construction and source of all variables, see Appendix A-1.

Table 11: Selective Survival in Baseline Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cohort:	> University	University	High School	Primary School	Pre School	Early Childhood	In Utero
<i>Panel A. Dependent variable is cell size divided by no. of years in cohort</i>							
Violence	-28.52* (15.90)	-35.46 (22.78)	-43.03 (30.39)	-50.22 (36.67)	-57.96 (41.76)	-77.80 (57.76)	-104.8 (76.16)
Constant	353.1* (185.8)	433.8 (277.6)	587.8 (379.3)	711.1 (487.7)	839.7 (575.9)	1279 (835.9)	1749 (1154)
R-squared	0.063	0.054	0.052	0.049	0.047	0.045	0.040
<i>Panel B. Dependent variable is share of boys in cohort</i>							
Violence	0.00337 (0.00448)	-0.000722 (0.00477)	-0.00155 (0.00480)	-0.00397 (0.00339)	0.00980** (0.00455)	3.71e-05 (0.00407)	-0.00407 (0.00526)
Constant	0.463*** (0.0298)	0.439*** (0.0334)	0.454*** (0.0273)	0.421*** (0.0233)	0.442*** (0.0331)	0.473*** (0.0264)	0.477*** (0.0338)
R-squared	0.061	0.024	0.019	0.039	0.025	0.023	0.056
Observations	372	372	372	372	372	372	372
Controls	✓	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29). Panel A regresses the size of each cohort (divided by the number of years per cohort) in each municipality on the violence dummy. Panel B regresses the share of boys in each cohort in each municipality on the violence dummy. All municipal controls are also included: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals' vote share in the 1946 presidential election (politics). For more details on the construction and source of all variables, see Appendix A-1.

Table 12: Mechanisms: Sex

	(1)	(2)	(3)	(4)
Dependent Variable	Years of Schooling		Mean Schooling in Job Sector	
Group:	Boys	Girls	Boys	Girls
Violence $\times$ In Utero	-0.352*** (0.127)	-0.179* (0.0913)	-0.168** (0.0827)	0.329** (0.150)
Violence $\times$ Early Childhood	-0.325*** (0.102)	-0.174** (0.0845)	-0.219** (0.0902)	0.222 (0.160)
Violence $\times$ Preschool	-0.279*** (0.0868)	-0.124* (0.0685)	-0.200** (0.0814)	0.117 (0.157)
Violence $\times$ Primary School	-0.172*** (0.0555)	-0.0669 (0.0474)	-0.134** (0.0574)	-0.00774 (0.136)
Violence $\times$ High School	-0.0935** (0.0471)	-0.0621 (0.0393)	-0.0703* (0.0395)	0.0928 (0.121)
Violence $\times$ University	-0.0953 (0.0904)	0.000546 (0.0403)	-0.0651 (0.0696)	0.0920 (0.108)
Controls $\times$ Cohort	Full	Full	Full	Full
Municipality Fixed Effects	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓
Observations	510,766	546,970	424,002	100,350
R-squared	0.359	0.334	0.433	0.284

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *Years of Schooling* represents the years of education of individuals (older than 6) in 1973. *Mean Schooling in Job Sector* is the average years of schooling of the employees in the sector on which the individual works. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. “Full” controls refers to the following municipal variables, which are interacted with cohort fixed effects: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals’ vote share in the 1946 presidential election (politics). For more details on the construction and source of all variables, see Appendix A-1.

**Table A-1: Identification of Municipality of Birth
1973 Census in Colombia**

	Individual was born in...					Total
	Known Municipality			Unknown Municipality		
	<i>Peaceful</i>	<i>Violent</i>	<i>Abroad</i>	<i>‘Comisarias’</i>	<i>Other</i>	
Total Individuals	15,371,221	2,491,612	138,876	120,289	276,388	18,398,386
Share of Total	83.55%	13.54%	0.75%	0.65%	1.5%	100%
Years of Schooling	3.23	2.97	3.72	2.31	2.85	3.19

Notes: *Years of Schooling* data are available for people older than 6 in 1973. *Comisarias* are the vast territories in the east and south of Colombia that were not formally constituted as departments in 1973; they were reported in the census as a single group, so the exact place of birth is unknown and it is impossible to assign the violence condition of their birthplace to this group.

Table A-2: Variables and Sources

Description		Source
A. Main Variables		
Years of Schooling	Years of formal education	Census 1973, National Administrative Department of Statistics
Mean Schooling in Job Sector	Average years of schooling of workers, by sector of employment and gender in 1973. Computed using only peaceful municipalities that are not geographical neighbors of violent municipalities. The sectors correspond to the International Standard Industrial Classification of all economic activities, Rev.2	Census 1973, National Administrative Department of Statistics
Migration	Dummy = 1 if municipality of birth is different from municipality of residence.	Census 1973, National Administrative Department of Statistics
Violence	Dummy = 1 if municipality of birth was highly affected by bipartisan violence from 1948 to 1953.	Revista Criminalidad Policía Nacional (1959)
B. Colonial Institutions		
Indigenous Population	Dummy = 1 if the municipality had an indigenous population from 1535 to 1540.	Fernández (2010)
Early Foundation Date	Foundation date of municipality (year).	Bernard and Zambrano (1993)
C. Geographical Covariates		
Log(Altitude of Municipality Head)	Log(altitude of municipality head above sea level, in meters)	CEDE, Universidad de los Andes
Rainfall (mm)	Mean annual rainfall level in millimeters in each municipality.	IDEAM, Institute of Hydrology, Meteorology and Environmental Studies
Roughness (Sd Altitude)	Standard deviation of altitude calculated for each municipality using the variance of its level curves in meters above sea level.	IGAC, Instituto Geográfico Agustín Codazzi: Colombia's national geographic institute.
D. Historical Conflict		
Conflict of Land Tenure (1901 to 1931)	Dummy = 1 for municipalities in which at least 5 families reported a conflict of land tenure with businessmen and landowners in the period between 1901 and 1931.	Fernández (2010) based on LeGrand (1988)
E. Politics		
Vote share for liberals in 1946	Votes for Liberals/ total votes in the 1946 presidential election	Registraduría Nacional del Estado Civil, Colombia. Presidential Elections 1946
F. Additional Robustness		
Bandoleros	Dummy = 1 if there were <i>bandoleros</i> from 1962 to 1965 in the municipality.	Revista Criminalidad Policía Nacional (1965)
Years of Late Violence	Number of years of exposure to violence after 1948.	Chacón, Robinson, and Torvik (2011)

Table A-3: Selective Survival in Full Sample of Municipalities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cohort:	> University	University	High School	Primary School	Pre School	Early Childhood	In Utero
<i>Panel A. Dependent variable is cell size divided by no. of years in cohort</i>							
Violence	-26.29** (11.33)	-28.40* (15.90)	-33.20 (20.83)	-31.62 (23.97)	-33.50 (26.42)	-45.90 (35.56)	-64.39 (45.05)
Constant	402.9*** (118.1)	512.5*** (173.0)	671.8*** (233.4)	792.3*** (293.7)	911.3*** (345.9)	1341*** (502.9)	1765** (688.6)
R-squared	0.042	0.035	0.035	0.033	0.031	0.031	0.029
<i>Panel B. Dependent variable is share of boys in cohort</i>							
Violence	0.00282 (0.00417)	-0.00313 (0.00416)	-0.00516 (0.00415)	-0.00254 (0.00312)	0.00626 (0.00406)	-0.000523 (0.00349)	-0.00729* (0.00441)
Constant	0.488*** (0.0204)	0.453*** (0.0216)	0.459*** (0.0188)	0.451*** (0.0167)	0.492*** (0.0215)	0.471*** (0.0183)	0.481*** (0.0237)
R-squared	0.015	0.026	0.033	0.022	0.019	0.021	0.022
Observations	732	732	732	732	732	732	732
Controls	✓	✓	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29). Panel A regresses the size of each cohort (divided by the number of years per cohort) in each municipality on the violence dummy. Panel B regresses the share of boys in each cohort in each municipality on the violence dummy. All municipal controls are also included: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals' vote share in the 1946 presidential election (politics). For more details on the construction and source of all variables, see Appendix A-1.

Table A-4: Mechanisms: Probability of No Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
Violence × Cohort in 1948	Dep. Variable: Dummy=1 if no schooling					
Violence × In Utero	-0.0389* (0.0222)	-0.00841 (0.0144)	-0.0287* (0.0163)	-0.0381* (0.0214)	-0.0378 (0.0257)	0.00660 (0.0159)
Violence × Early Childhood	-0.0302* (0.0180)	-0.00872 (0.0124)	-0.0253* (0.0145)	-0.0301* (0.0179)	-0.0281 (0.0213)	0.00391 (0.0142)
Violence × Pre School	-0.00685 (0.0138)	0.00884 (0.0103)	-0.00549 (0.0120)	-0.00705 (0.0140)	-0.00769 (0.0162)	0.0140 (0.0121)
Violence × Primary School	-0.00711 (0.0114)	0.00555 (0.00842)	-0.00739 (0.0102)	-0.00739 (0.0116)	-0.00837 (0.0134)	0.00921 (0.00999)
Violence × High School	0.000239 (0.00698)	0.00386 (0.00637)	-0.00243 (0.00683)	6.28e-05 (0.00715)	0.00198 (0.00857)	0.00883 (0.00717)
Violence × University	-0.00483 (0.00604)	-0.00183 (0.00552)	-0.00555 (0.00607)	-0.00488 (0.00611)	-0.00489 (0.00699)	0.000244 (0.00620)
Sex	0.0209*** (0.00307)	0.0210*** (0.00307)	0.0210*** (0.00309)	0.0210*** (0.00307)	0.0183*** (0.00297)	0.0183*** (0.00299)
Constant	0.290*** (0.00693)	0.289*** (0.00515)	0.290*** (0.00652)	0.290*** (0.00701)	0.282*** (0.00750)	0.280*** (0.00594)
Controls × Cohort	No	Institutions	Geography	H.Conflict	Politics	Full
Municipality Fixed Effects	✓	✓	✓	✓	✓	✓
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Observations	1,162,192	1,162,192	1,157,696	1,162,192	1,099,801	1095305
R-squared	0.161	0.161	0.161	0.161	0.161	0.162

Note: Robust standard errors in parentheses clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1. Cohorts defined by age in 1948: prebirth/in utero individuals (1-2 years before birth), early childhood (0 to 3), preschool (4 to 6), primary school (7 to 12), high school (13 to 17), university (18 to 23), and beyond university (24 to 29), which acts as the excluded, comparison cohort. The set of municipal controls interacted with cohort fixed effects are marked in each column and include the following: presence of indigenous population from 1535 to 1540 and foundation date of the municipality (institutions), altitude of municipality head, rainfall and roughness (geography), conflicts of land tenure from 1901 to 1931 (historical conflict), Liberals' vote share in the 1946 presidential election (politics). "Full" refers to the entire set of controls from Columns 2 to 5. For more details on the construction and source of all variables, see Appendix A-1.