

Online-Appendix

A *La Violencia*: General Context

Long-standing political power struggles between the two traditional parties (Liberal and Conservative) that had turned violent numerous times during the XIXth Century and most notably during the *War of the Thousand Days* from 1899 to 1902, revived in 1948 after the assassination of the liberal presidential candidate, Jorge Eliécer Gaitán, leading to the period commonly referred to as *La Violencia*. The most reliable 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 (out of a population of around 11.5 million). 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](#)).

These national-level political struggles fueled long-standing local disputes for political control and land access, as well as revenge raids and banditry ([Oquist, 1980](#); [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 et al. \(1964\)](#) estimate that 200,000 were killed up to 1962. As a percentage of the population of 11 million, these casualties are higher than the sum of all American deaths during the two world wars and the Korean wars combined ([Lipman & Havens, 1965](#)). Violence was also particularly horrid. It was common to disfigure or cut the dead into pieces, with some cutting procedures even

getting specific names (see, for example, [Uribe \(2004\)](#)).

In addition to deaths and maiming, the violence also caused several economic losses. Due to forced migration and illegal land seizure, a little less than 394,000 hectares of land were lost ([Oquist, 1980](#)). 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, especially in rural areas: 41% and 78% of urban and rural teachers, respectively, had primary education or less ([Havens, 1965](#)).

As for the geographic dispersion of violence in the Colombian territory, the literature has few definite explanations for the observed variation. On the one hand, a bulk of the studies of *La Violencia* are regional examinations of the social underpinnings and repercussions of violence (e.g. [Arocha, 1979](#); [Sánchez & Meertens, 1983](#); [Ortiz, 1985](#); [Henderson, 1985](#); [Betancourt & García, 1990](#); [Roldan, 2002](#)), offering diverse explanations but without

a systematic empirical examination at the national level. On the other hand, research with a national scope has focused on interpreting *La Violencia* broadly as a consequence of the features of the Colombian state, in particular its weakness and failure to legitimately channel democratic demands (Oquist, 1980; Pécaut, 1987). An exception is the work of Chacón, Robinson, and Torvik (2011), which empirically examines the geographical dispersion of violence and shows that it was particularly prevalent in municipalities where political competition was highest. In any case, it is clear that violence did not occur randomly. However, given the results of both the falsification tests and the robustness specifications allowing for differential trends in education as a function of socioeconomic and political features of municipalities, as we argue in the main text we can be confident about the validity of our identification strategy to uncover the effects of violence.

B Data sources and variables

As noted in the main text, we measure violence as a simple dummy variable indicating *La Violencia* incidence as reported by the National Police on *Revista Criminalidad*. This source has been used in several studies of *La Violencia*, like Chacón et al. (2011) and Chacón (2004). Unfortunately, the report only provides a measure of the extensive margin, detailing which municipalities were affected by violent events during the period of *La Violencia*. However, it has two key advantages. First, with reports published in 1958 and 1959 after the violence and Rojas' dictatorship, and during the initial years of the the National Front, we have no reason to suspect systematic partisan biases in the coding of violence. In fact, the police reports is not only independent but also highly critical of political parties in its description of *La Violencia*. Second, this was a truly national effort, including the entire territory of Colombia. Intensive margin measures (approximate number of deaths, years affected by violence) are available in the literature, but they are highly speculative numbers, or available only for regions of the country, or both. Indeed, total estimated deaths in existing

historical sources range from 200,000 to 400,000 (see for example [Roldan \(2002\)](#) for a discussion). We thus consider our binary measure less likely to be influenced by severe measurement errors. In line with this, where available, the incidence dummy matches information from more detailed regional studies of this episode of civil conflict ([Oquist, 1980](#); [Roldan, 2002](#); [Fernández, 2010](#)).

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).

Table B-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.

**Table B-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 B-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) |
| Teacher | Dummy = 1 if sector of employment is equal to 931 or 932 of the CIU classification revision 2. These codes correspond to the sectors of Education Services and Research and Scientific Institutes. | Census 1973, National Administrative Department of Statistics |
| Sector of employment Definition 1 | Variable that takes values from 1 to 3 according to the sector of employment using the CIU classification revision 2. 1=agriculture, 2=manufacturing, 3=services. It uses the following sub sectors as categories: 1) Agriculture, Hunting, Forestry and Fishing, 2) Manufacturing and 3) Financing, Insurance, Real Estate and Business Services and missing otherwise. | Census 1973, National Administrative Department of Statistics |
| Sector of employment Definition 2 | Variable that takes values from 1 to 3 according to the sector of employment using the CIU classification revision 2. 1=agriculture, 2=manufacturing, 3=services. It uses: 1) Agriculture, Hunting, Forestry and Fishing 2) Manufacturing, Construction, Electricity, Gas and Water, 3) Financing, Insurance, Real Estate and Business Services and missing otherwise. | Census 1973, National Administrative Department of Statistics |
| 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) |

C Additional empirical exercises and robustness

C.1 Results by migration status

See Table C-1.

C.2 Yearly cohorts

See Figures C-1 and C-2.

C.3 Violence persistence

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 C-2 and C-3 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 (*Bandoleros*) in the early 1960s (1962 and 1965) 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. Thus, they must be interpreted with caution (see Figure C-4 for a timeline of the events). However, if violence persistence explains our findings, then including these bad controls should reduce the size of our estimated effects and 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 C-2

with our baseline results for years of schooling (Column 1 of Table 2) 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 2. Table C-3 runs analogous regressions for the average schooling in the job sector as the dependent variable, and can be compared to results in Table 4. The finding is similar: a minor impact on the size and significance of our main treatment interactions.²⁶

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.

C.4 Department differential trends

Some economic policies, infrastructure investments and educational programs during the 50s and 60s in Colombia were decided and implemented at department level (Ramirez & Tellez, 2006). To check whether our results are driven by specific trends related to these policies we re-estimate our baseline results allowing for differential trends at this level. The identification in this case comes from the variation in our outcome variables across municipalities within the same department, and not from comparisons of municipalities across departments. Tables C-4 and C-5 present the results of this exercise using our two outcome variables. In both tables we still find a negative and sizeable effect of *La Violencia* on years of schooling and schooling in job sector.

²⁶To save space in the tables, we omit the coefficients for the interactions of measures of persistence and cohorts – which are typically not significant.

C.5 Sample selection: neighboring municipalities

In our main regression tables, we focused on neighboring municipalities to study the impact of violence. In this section we discuss the implications of restricting the sample on this dimension.

Recall the two main reasons that motivated our decision to focus on them 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 sector-of-employment regressions. Indeed, it enables us to estimate the typical human capital content of each employment sector 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.

At the same time, this last observation also raises the question of whether our peaceful or ‘control’ municipalities are affected by spillovers from their violent neighbors. For instance, 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. To study this possibility, Column 1 in Table C-6 verifies 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 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 municipalities.

Spillovers from migration, however, may be present even when the overall region is not directly affected by violence. These spillovers may be positive (increasing average educational attainment, thus leading us to overestimate the costs of violence) or negative (with the opposite implication). When 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. 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.

To investigate these issues, Column 2 in Table C-6 focuses on migrants only, and explores 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. Figure 4 already suggests that the cohort composition of migrants from violent municipalities is similar to those from their peaceful neighbors. But this does not fully rule out all these possible spillovers. In Panel B (Column 3) of Table C-6, 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.

These results 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 into control areas. However, the results are only suggestive and disentangling each of these possible influences is very challenging and not feasible with the available data. As a final exercise, 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 C-3, 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 C-1 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.

**Figure C-1: Difference-in-Differences Estimate of the Effect of Violence on Schooling
By Age in 1948**



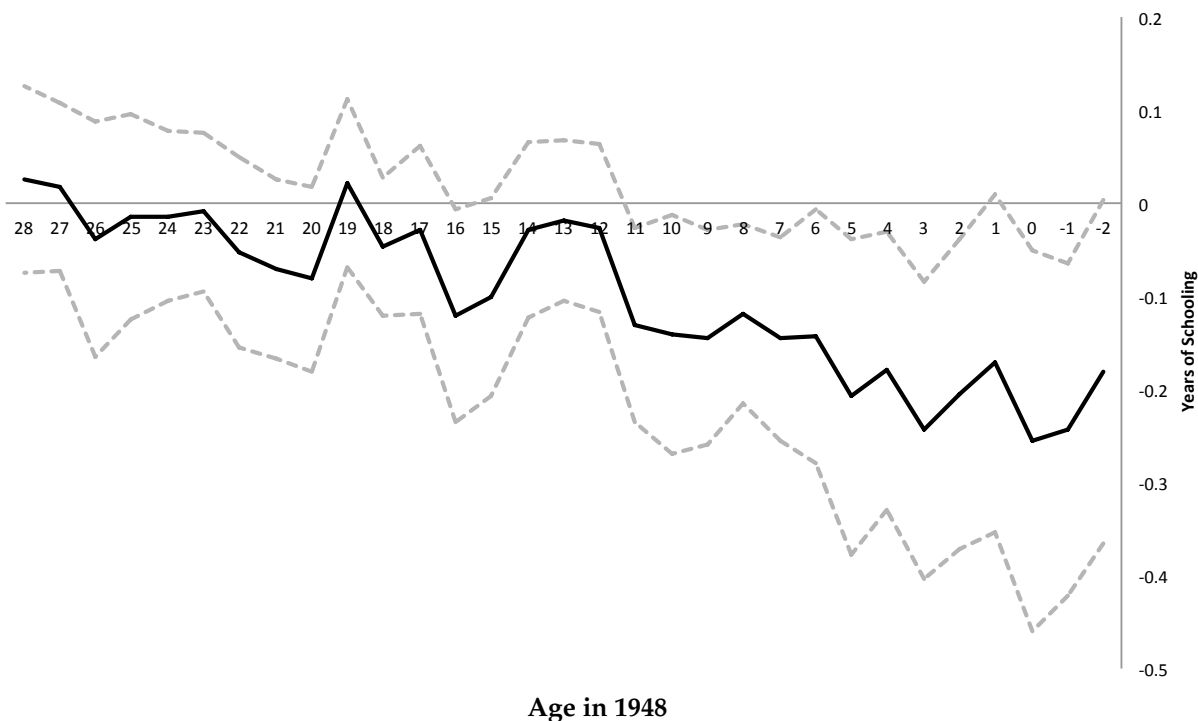
Note: The figure depicts the difference-in-differences estimates of the impact of violence on years of schooling for each cohort, by age in 1948. It plots coefficients and 90% confidence intervals for γ_k in the following regression model: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{imk}) \gamma_k + \epsilon_{imk}$, 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_{imk} 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_{imk} set is that aged 29).

Figure C-2: 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: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{imk}) \gamma_k + \epsilon_{imk}$, 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_{imk} 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_{imk} 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 C-3: 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: $y_{imk} = c + \alpha_m + \lambda_k + \sum_k (V_m \times d_{imk}) \gamma_k + \epsilon_{imk}$, 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_{imk} 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_{imk} is that aged 29).

Figure C-4: Sequence of Events
La Violencia in Colombia

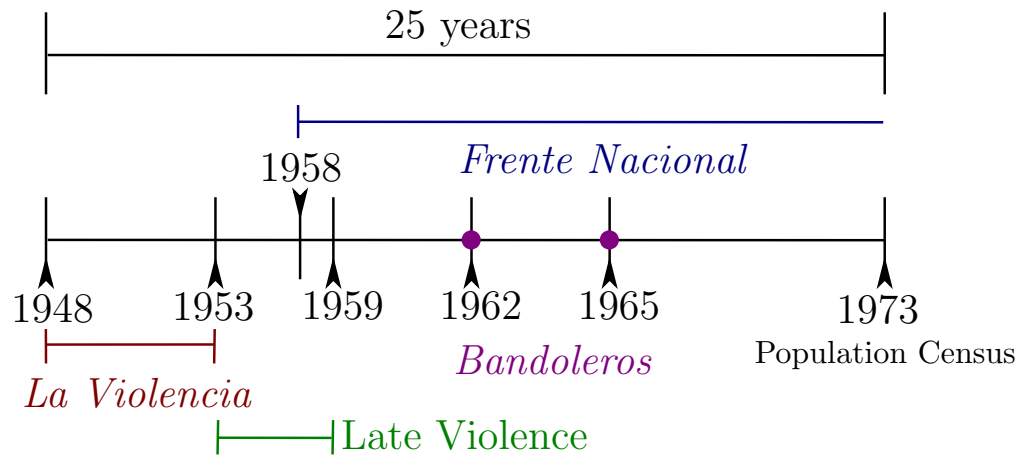


Table C-1: Effects of *La Violencia* on Schooling and Mean Schooling in Job Sector By Migration Status

| Dependent Variable | Years of Schooling | | | Mean Schooling in Job Sector | | |
|-----------------------------------|-----------------------|-----------------------|-----------------------|------------------------------|----------------------|-----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Violence × Cohort in 1948 | | | | | | |
| 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 1 (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 C-2: Persistence of Violence: Years of Schooling

| | (1) | (2) | (3) | (4) |
|--|--|-----------------------|-----------------------|-----------------------|
| Violence × Cohort in 1948 | Dep. Variable: Years of Schooling | | | |
| Violence × In Utero | -0.291*** (0.103) | -0.235** (0.0962) | -0.243** (0.0987) | -0.253** (0.0993) |
| Violence × Early Childhood | -0.288*** (0.106) | -0.207** (0.0826) | -0.213** (0.0899) | -0.210** (0.0825) |
| Violence × Pre School | -0.249*** (0.0879) | -0.169** (0.0655) | -0.167** (0.0657) | -0.152** (0.0623) |
| Violence × Primary School | -0.103*** (0.0394) | -0.0872** (0.0402) | -0.0861** (0.0396) | -0.0903** (0.0400) |
| Violence × High School | -0.0585* (0.0306) | -0.0653** (0.0321) | -0.0716** (0.0316) | -0.0774** (0.0314) |
| Violence × 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 × Cohort | ✓ | ✓ | - | - |
| Years of Late Violence × Cohort | - | - | ✓ | ✓ |
| Full Controls × 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 C-3: Persistence of Violence: Structural Change

| | (1) | (2) | (3) | (4) |
|--|--|----------------------|----------------------|----------------------|
| Violence × Cohort in 1948 | Dep. Variable: Mean Schooling in Job Sector | | | |
| Violence × In Utero | -0.214** (0.0971) | -0.137 (0.0867) | -0.137 (0.0885) | -0.119 (0.0836) |
| Violence × Early Childhood | -0.301** (0.116) | -0.185** (0.0933) | -0.214** (0.102) | -0.170* (0.0893) |
| Violence × Pre School | -0.289*** (0.104) | -0.157** (0.0769) | -0.186** (0.0829) | -0.122* (0.0716) |
| Violence × Primary School | -0.189*** (0.0701) | -0.117** (0.0531) | -0.136** (0.0586) | -0.102** (0.0492) |
| Violence × High School | -0.0924** (0.0413) | -0.0479 (0.0387) | -0.0698* (0.0397) | -0.0376 (0.0367) |
| Violence × 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 × Cohort_k | ✓ | ✓ | - | - |
| Years of Late Violence × Cohort_k | - | - | ✓ | ✓ |
| Full Controls × 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 C-4: Robustness: Years of Schooling - Department Trends

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------------------------|--|-----------------------|-----------------------|-----------------------|-----------------------|----------------------|
| Violence × Cohort in 1948 | Dep. Variable: Years of Schooling | | | | | |
| Violence × In Utero | -0.167** (0.0662) | -0.123** (0.0613) | -0.149** (0.0612) | -0.168*** (0.0608) | -0.231*** (0.0747) | -0.141** (0.0617) |
| Violence × Early Childhood Age | -0.157*** (0.0595) | -0.121** (0.0537) | -0.140** (0.0570) | -0.157*** (0.0575) | -0.216*** (0.0665) | -0.128** (0.0536) |
| Violence × Pre School Age | -0.143*** (0.0510) | -0.129*** (0.0497) | -0.128** (0.0502) | -0.143*** (0.0503) | -0.161*** (0.0570) | -0.107** (0.0518) |
| Violence × Primary School Age | -0.0848** (0.0334) | -0.0832** (0.0343) | -0.0731** (0.0335) | -0.0841** (0.0335) | -0.0822** (0.0359) | -0.0582 (0.0362) |
| Violence × High School Age | -0.0489* (0.0282) | -0.0451 (0.0292) | -0.0398 (0.0282) | -0.0485* (0.0281) | -0.0645** (0.0299) | -0.0455 (0.0311) |
| Violence × University Age | -0.0359 (0.0236) | -0.0368 (0.0244) | -0.0294 (0.0238) | -0.0357 (0.0235) | -0.0346 (0.0257) | -0.0216 (0.0272) |
| Sex | -0.351*** (0.131) | -0.351*** (0.131) | -0.352*** (0.131) | -0.351*** (0.131) | -0.359*** (0.137) | -0.360*** (0.137) |
| Controls × Cohort | No | Institutions | Geography | H.Conflict | Politics | Full |
| Department Dummies × Cohort | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 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 of controls from Columns 2 to 5. For more details on the construction and source of all variables, see Appendix A-1.

Table C-5: Robustness: Structural Change - Department Trends

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------------------------|---|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|
| Violence × Cohort in 1948 | Dep. Variable: Mean Years of Schooling in Job Sector | | | | | |
| Violence × In Utero | -0.144** (0.0671) | -0.0987* (0.0583) | -0.125* (0.0663) | -0.145** (0.0630) | -0.192** (0.0754) | -0.111* (0.0634) |
| Violence × Early Childhood Age | -0.179** (0.0725) | -0.129** (0.0612) | -0.168** (0.0729) | -0.180*** (0.0695) | -0.240*** (0.0816) | -0.156** (0.0662) |
| Violence × Pre School Age | -0.146** (0.0620) | -0.116** (0.0559) | -0.121* (0.0624) | -0.147** (0.0603) | -0.179** (0.0698) | -0.112* (0.0602) |
| Violence × Primary School Age | -0.128*** (0.0443) | -0.107*** (0.0401) | -0.109** (0.0449) | -0.129*** (0.0423) | -0.148*** (0.0498) | -0.0945** (0.0420) |
| Violence × High School Age | -0.0588* (0.0350) | -0.0481 (0.0356) | -0.0489 (0.0363) | -0.0596* (0.0336) | -0.0575 (0.0386) | -0.0235 (0.0380) |
| Violence × University Age | -0.0168 (0.0248) | -0.0166 (0.0246) | -0.00744 (0.0251) | -0.0176 (0.0239) | -0.0179 (0.0278) | -0.00589 (0.0274) |
| Sex | 0.945*** (0.151) | 0.944*** (0.151) | 0.946*** (0.152) | 0.945*** (0.151) | 0.934*** (0.153) | 0.935*** (0.153) |
| Constant | 3.214*** (0.0243) | 3.214*** (0.0246) | 3.215*** (0.0239) | 3.215*** (0.0228) | 3.277*** (0.0251) | 3.277*** (0.0237) |
| Controls × Cohort | No | Institutions | Geography | H.Conflict | Politics | Full |
| Department Dummies × Cohort | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 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 of controls from Columns 2 to 5. For more details on the construction and source of all variables, see Appendix A-1.

Table C-6: Violence, Migration and Education

| Panel A: Violence and Migration | | | Panel B: Migration and Education | | |
|---|----------------------|------------------------|-----------------------------------|-----------------------------|----------------------|
| | (1) | (2) | | (3) | (4) |
| 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