Violence and Ethnic Identity: Evidence from Peru*

Aldo Elizalde
Queen's University Belfast†

Eduardo Hidalgo
University of Cologne‡

Sotiris Kampanelis
Cardiff University§

March 6, 2025

Abstract

This paper examines the consequences of violence on ethnic identity using the Shining Path conflict in Peru as a natural experiment. The Shining Path conflict was among the most brutal civil wars in 20th-century Latin America—claiming over 70,000 lives, subjecting many to torture, rape, and displacement. The insurgency arose from an extremist revolutionary ideology in rural areas that demanded class identity over any other identity affiliation, causing Indigenous populations to bear 75% of the victims. Leveraging individual-level ethnic identity data and violent event-level data from 1958 to 1992, we implement a difference-in-differences strategy that exploits quasi-random variation in exposure to violence across individuals at different stages of their ethnic identity formation. Our findings indicate that individuals exposed to conflict violence during their crucial formative years of ethnic identity are significantly less likely to identify as Indigenous or to speak an Indigenous language as their mother tongue compared to those who were not exposed and whose ethnic identity has been largely established. We attribute these results to own-ethnic group violence, which weakens incentives to maintain Indigenous identity.

Keywords: Ethnic identity; Violence; Civil Conflict; Indigenous Populations, Peru

JEL Codes: D74; J15; O15; N36;

*We thank Gustav Agneman, Miriam Artiles, Francisco Eslava, Eric Melander, Alberto Hidalgo, Erik Hornung, Juan Pablo Posada, Susanne Prantl, Carola Stapper for helpful comments.

†Department of Economics, Queen's University Belfast; a.elizalde@qub.ac.uk

‡University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany; hidalgo@wiso.uni-koeln.de

§Cardiff Business School, Cardiff University; kampaneliss@cardiff.ac.uk

1. Introduction

One of the most salient identities for human beings is their ethnic identity, which defines who they are and their behavior (Akerlof and Kranton, 2010, p. 13, 28). While the existing economic literature has found a clear impact of ethnic diversity on violence conflict, there is a lack of understanding on how violence affects ethnic identity (Alesina et al., 1999; Arbatlı et al., 2020; Esteban et al., 2012; Montalvo and Reynal-Querol, 2005; Rohner et al., 2013a,b).

In this paper, we investigate the consequences of violence on ethnic identity by using the Shining Path conflict in Peru as a natural experiment. At the beginning of the 1980s, Peru's Shining Path guerrilla started its revolutionary activity by settling in rural areas and pressuring the population to embrace class identity. In this conflict, more than 70.000 people were killed, and many more suffered torture, rape, and displacement. The Indigenous population was the most affected ethnic group representing 75% of the victims of this conflict while only comprising 30% of the total population.

This setting provides two advantages for studying the consequences of violence on ethnic identity. First, while the Shining Path conflict did not target Indigenous people *per se*, they were their main victims. The Shining Path ideology emphasized that any other identity but the class identity, particularly peasantry, was alienation from capitalism and enforced the abandonment of those other identities (Jima-González and Paradela-López, 2020). In this context, all people, including the Indigenous people, had the decision to abandon their ethnic identity to embrace the class identity or suffer potentially violent consequences. As one of the insurgency's leaders put it: *"First they [rural population] will be given political reeducation, and if that fails, they will receive what we get now: dictatorship, prison ... death. The genocide which we have accepted, they will have to accept"* (Wheat, 1990, p. 48). Second, the Shining Path conflict lasted from 1982 to 1992 and affected approximately one-third of Peruvian districts. Therefore, this conflict provides a large source of variation at the time and spatial level, which allows us to exploit differences in exposure to violence between cohorts of individuals over a ten year period across one-third of Peruvian districts.

To study the consequences of violence on ethnic identity, we combine individual-level data on ethnic identity and violent event-level data from 1958 until 1992. Specifically, we collect individual-level data on ethnic identity from the *Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza* (ENAHO). We measure ethnic identity in two ways. First, based on self-identification, an individual is classified as Indigenous if individual i identifies itself as Quechua, Aymara or another Indigenous group, and we classify it as non-Indigenous if individual i identifies as Mestizo. Second, based on mother tongue, an Individual is classi-

fied as Indigenous if their mother tongue is Quechua, Aymara or another Indigenous language, and they are classified as non-Indigenous if Spanish (Castellano) is their mother tongue.

We collect data on violent events from the Peruvian Truth and Reconciliation Commission (TRC) database. This dataset is the most comprehensive dataset on the Shining Path conflict. It provides event-level data on the location, the time, the type of violence, and the ethnic identity of the victim of the violent event for more than 36,000 unique violent events from 1980 until 2000. We combine the TRC dataset with the ENAHO dataset to create an exposure to the violence variable at the individual level. The idea of the identification strategy is to compare ethnic identity differences of individuals that were exposed and not exposed to violence during the years when their identity is formed (0-19) with respect to ethnic identity differences of individuals that were exposed and not exposed to violence when identity was already formed ($>$19). Last, since we are interested in assessing the impact of conflict on Indigenous identity, we only consider violent events where an Indigenous person is a victim.

The main concern of the identification strategy is that violent events did not occur randomly, and the Shining Path could have targeted some specific districts based on the identity characteristic of the population. To address this issue, we use a border sample design that allows us to select a set of districts that are equal in socio-economic characteristics but suffered different levels of violence.

We implement a difference-in-differences strategy that exploits quasi-random variation levels of exposure to violence across individuals using the bordering sample districts. This allows us to compare ethnic identity differences of individuals exposed and not exposed to violence during the ethnic identity formation years (0-19) compared to individuals exposed and not exposed to violence when ethnic identity was already formed ($>$19). To further understand the consequences of violence on identity, we use a cohort-level analysis and compare the probability of individuals identifying as Indigenous conditionally on the year-cohort $\in$ (0-4, 5-9, 10-14, 15-19) that they were exposed to violence compared to the cohort (20-24) when identity was supposed to be formed at the time of violence.

The estimates show that individuals exposed to violence are 2.3 percentage points less likely to identify as Indigenous and 3.4 percentage points less likely to speak an Indigenous language as their mother tongue than comparable individuals that were not exposed to violence. These results are meaningful in size since a 2.3 (3.4) percentage points decrease in the probability of identifying as Indigenous (speak an indigenous language as mother tongue) represents a 5.75% (13.6%) decrease with respect to the mean (40%). Furthermore, the results of the cohort-level analysis show that individuals exposed to violence in

their childhood were the most affected. The estimates suggest that individuals who experienced violence between 0 and 4 years old are 7.8 percentage points less likely to speak an Indigenous language as their mother tongue and 4.7 percentage points less likely to identify Indigenous than individuals who did not experienced violence in similar cohorts. Likewise, individuals who experienced violence between 5 and 9 years old are 5.3 percentage points less likely to speak an indigenous language as their mother tongue and 3.2 percentage points less likely to identify Indigenous than individuals who did not experienced violence in similar cohorts. We do not find statistical differences in the probability of speaking an Indigenous language as a mother tongue or identifying as Indigenous between the 10-14 and 15-19 cohorts and the 20-24 cohort.

The validity of the identification strategy relies on the assumption that in the absence of violence, Indigenous identity would have followed common trends between cohorts of individuals exposed and not exposed to violence. To provide evidence for this assumption, we compare Indigenous identity differences between individuals that were exposed to violence and not exposed to violence when their ethnic identity was already formed (cohorts 25-29, and 30-34) with respect to our baseline cohort (20-24). We find no statistical differences in ethnic identity between these cohorts. Further, we compare Indigenous identity of individuals born in districts that experienced violence after violence occurred and the baseline cohort (20-24). The results show no difference between these two groups in their likelihood of identifying as Indigenous or speaking an Indigenous language as their mother tongue.

We conduct additional robustness checks to validate the main results. First, we test whether exposure to non-Indigenous violence (when the victim is not an Indigenous person) impacts ethnic identity. The estimates show that exposure to non-Indigenous violence has no consequences on the probability of an individual identifying as Indigenous. Second, we check whether migration is confounding the main results. When controlling for migration, the results remain similar, and when using migration as an outcome, we do not find a significant effect of exposure to violence on migration. Third, we check whether exposure to violence had a consequence on class and national identity. We find no effect on class-identity and a negative association between exposure to violence and a national identity feeling. Last, we test whether exposure to violence affected the different Indigenous groups differently. The results show that violence mostly affected Indigenous identity of the Quechua population and Indigenous groups that were not Aymara.

To understand the mechanism driving the main results, we build on the recent advances in understanding the dynamics of violent conflicts, trust, and ethnic identity. Rohner et al. (2013a,b) argue that

exposure to violence in a conflict erodes generalized trust in a society and enhances interpersonal trust and ethnic identification. The main difference between Rohner et al. (2013a,b) and the Peruvian case is that in the former case violence was perpetuated between different ethnic groups, whereas in the Shining Path conflict in Peru it was perpetuated locally by people living in the same village. Therefore, we expect the conflict to erode interpersonal trust and Indigenous identity in Indigenous homogeneous regions.

To test the consequences of exposure to violence on trust, we collect data on interpersonal trust from the LAPOP surveys, generalized trust from the Latinobarometro surveys, and institutional trust from the ENAHO surveys. The results show that individuals exposed to violence have lower interpersonal trust, and (weakly) higher generalized trust than comparable individuals that were not exposed to violence. Further, the results show that trust in different state institutions is not different between individuals that were exposed to violence and individuals that were not exposed to violence. We hypothesize that these results are opposite to Rohner et al. (2013a,b) because individuals were exposed to violence by their own Indigenous group.

To test whether own-Indigenous group violence is the mechanism driving the main results, we construct a measure of Indigenous diversity at the district level using ethnic identity data of individuals born before 1950 following Alesina et al. (1999). This index measures the probability that two randomly drawn people from the same district belong to different Indigenous groups. The rationale for using these measures is that in more Indigenous homogeneous districts, the likelihood of suffering violence by your own ethnic group is higher, leading to a lower Indigenous identity. To test this hypothesis, we interact the Indigenous diversity index with the exposure to violence variable. The results show that in Indigenous homogeneous districts, individuals exposed to violence are less likely to speak an Indigenous language as a mother tongue and to identify as Indigenous than comparable individuals not exposed to violence. Furthermore, a 1-sd increase in ethnic diversity index increases the probability of speaking an Indigenous language by 5.8 percentage points and increases the probability of identifying as Indigenous by 1.1 percentage point.

This paper contributes to three branches of the economic literature. First, it contributes to the growing literature on economics and identity.[1] This literature provides theoretical and empirical evidence about how identity, or a person's sense of self, affects economic outcomes (Akerlof and Kranton, 2000). Chen and Li (2009) provide evidence that

[1] see: Akerlof and Kranton, 2000, 2002; Aspachs-Bracons et al., 2008; Bénabou and Tirole, 2011; Chen and Li, 2009; Clots-Figueras and Masella, 2013; Constant and Zimmermann, 2008; Hargreaves Heap and Zizzo, 2009; Kranton, 2016; Schüller, 2015; Xin Li, 2010

when individuals share identity, they are more prone to within-group charity and lower envy. Furthermore, individuals are more likely to choose social-welfare-maximizing actions when sharing a group's identity. Xin Li (2010) shows that ethnic and national identities shape the tax morale of a country. The author finds that ethnically fractionalized countries have poorer tax morale than ethnically homogenous countries. In a lab experiment, Hargreaves Heap and Zizzo (2009) measure the social value of groups. The authors find that trust falls due to negative discrimination against outsiders when including outsiders in a group.

However, the economic determinants of ethnic identity have remained largely unexplored. Clots-Figueras and Masella (2013) provide evidence of education policies on shaping ethnic identity. Using the change of education policy by introducing Catalan along side with Spanish in Catolonia, the authors show that individuals that were exposed to Catalan have stronger Catalan feelings. Battu and Zenou (2010) provide evidence that the social environment and attachment to culture of origins are determinants of ethnic identity. Additionally, the study finds that ethnic identity has an impact on labour market success. Aspachs-Bracons et al. (2008) provide evidence that mandatory language policies have an effect on ethnic identity, whereas policies in which parents can choose the language have no impact on ethnic identity. We contribute to this literature by providing empirical evidence that exposure to violence has a negative association with ethnic identity.

Second, this paper contributes to the established literature linking ethnicity and conflict.[2] This literature shows, theoretically and empirically, how different measures of ethnic diversity correlate with civil conflict. For example, seminal work by Alesina et al. (1999) argues that the underprovision of public goods is related to ethnic heterogeneity, and ethnic conflict is an important determinant of public finances. Montalvo and Reynal-Querol (2005) provide empirical evidence that ethnic polarization is an important determinant of civil wars, and La Porta et al. (1999) provide evidence that ethnic heterogeneity is correlated with low efficiency in governments and corruption. In a series of articles Esteban et al. (2012) and Esteban and Ray (2011) study the link between ethnicity and conflict. First, the authors develop a theoretical model to understand the link between ethnic polarization and ethnic fractionalization with conflict, and second, they provide empirical evidence for their theoretical predictions. They show that ethnic polarization, fractionalization, and the Gini-Greenberg index are significantly correlated with conflict.

[2] see: Collier and Hoeffler (2004), Collier et al. (2009); Montalvo and Reynal-Querol (2005); Esteban et al. (2012), Esteban and Ray (2011), Rohner (2011), Caselli and Coleman (2013)

However, our understanding of the effect of violent conflict on ethnic identity is still limited. Rohner et al. (2013b) develop a theory of persistent ethnic conflicts. The authors argue that conflict lead to a collapse of trust, fuelling more ethnic conflict. Rohner et al. (2013a) provide empirical evidence for Rohner et al. (2013b) theoretical predictions. In Rohner et al. (2013a), the authors study the impact of social conflict on trust and ethnic identity. They find that more intense conflict is associated with lower generalized trust and higher ethnic identification. This paper complements this literature by providing empirical evidence regarding the heterogeneous effects of exposure to violence on ethnic identity depending on whether violence was perpetrated by a different ethnic group or by your own ethnic group.

Last, we contribute to the growing literature studying the causes and consequences of the Shining Path guerrilla. Albertus (2020) studies the determinants for the Shining Path uprising. The author provides evidence that greater land reform dampened Shining Path conflict events. García-Ponce (2017) investigates the association between Shining Path insurgency and women's engagement in local politics. The author provides evidence that gender quotas in Peru were more successful in municipalities exposed to the uprising. Galdo (2013) studies the consequences of Shining path conflict on labor market outcomes. The study finds that exposure to violence during the first 36 months of life is most detrimental to future labor-market outcomes. The author finds that a 1-sd increase in the level of violence exposure during the first 36 months of life is associated with 5% lower monthly earnings, a 3.5% lower probability of working in the formal sector and a 6% lower probability of working in large firms. Leon (2012) studies the consequences of the Shining Path conflict on human capital accumulation. The author finds that individuals exposed to conflict accumulate less years of education, and that these results are stronger in the short-term than in the long-run. We contribute to this literature by proving evidence that the Shining path conflict not only affected economic outcomes but changed individuals' identities.

2. Background

2.1. The Shining Path conflict

On May 17th, 1982, in the locality of Chuschi in the Department of Ayacucho, a group of 5 hooded men broke into premises where the electoral boxes and ballots were kept and burnt them. This event was claimed by the Shining Path and marked the beginning of the civil conflict in Peru. The confrontation between the Shining Path (Sendero Luminoso or Sendero) guerrilla and the Peruvian state was one of the most brutal civil conflicts endured throughout Latin America during

the 20th century. During this conflict, 70,000 people were estimated to be killed, and many more suffered torture, rape, and displacement. The conflict mostly affected the poor rural population, especially the Indigenous population (CVR, 2003).

The Shining Path organization results from an ideological conflict within the Peruvian Communist Party (PCP). Abimael Guzman, the leader of the Shining Path organization, advocated for a violent revolution to achieve socialism in contrast to the majority of leaders of the PCP. Guzman saw the Maoist *"Cultural Revolution"* as the most remarkable political event humanity has ever seen (CVR, 2003, p. 15). The division between the Shining Path and PCP occurred in 1970, and Abimael Guzman lost the ideological fight. At the time of the rupture, the Shining Path had 12 members in Ayacucho and 50 in the rest of the country (CVR, 2003, p. 17).

During the 1970s, Guzman's organization was primarily based in the Department of Ayacucho, where he was a professor at the University of Humanga. In this decade, the organization members were university professors, university students and rural teachers, who held clandestine meetings to re-construct the PCP with Guzman's ideas. Following other communist parties, they constructed the Shining Path from top to bottom and from ideology to politics. Furthermore, this group of professors and students would be at the top of the hierarchy that was supposed to bring the peasantry to the revolutionary success (CVR, 2003, p. 17 -19).

Based on their study and learnings from Marxism, the main Shining Path leaders started the diffusion of their ideology. They started the diffusion of their ideology first to university students who would eventually graduate as teachers and would be able to diffuse the Shining Path ideology through secondary schools. In this way, the Shining Path created a small but ideologically cohesive groups dependent on the party. The idea was that once the armed conflict started, the peasantry would join the Shining Path guerrilla under the party's direction.

May 17th, 1982, marked the beginning of the armed conflict. In this year, the Shining Path was present in the rural areas of Ayacucho, and its discourse of social justice and equality was accepted primarily by poor farmers. From 1983, the Shining Path started a more coercive campaign to dominate the peasantry. This increased the death toll of people that were against the revolution, community leaders and "richer" peasants that the Shining Path identified as enemies of the people (CVR, 2003, p. 43-46).

The Shining Path rhetoric was caught on the young population, mainly individuals between 12 and 30 years old. This rhetoric was appealing because of the promise of a "new world order" that would bring justice and social equality. These new ideas were an earthquake to the Andean cultural traditions, especially for the Indigenous people,

because in their traditional structures, power was held by the older population, and all the population respected this. This "new order" changed the hierarchical structures by giving the power to children and young adults, while the rest of the population lived in fear of losing their lives if they did not agree with this order (CVR, 2003, p. 47).

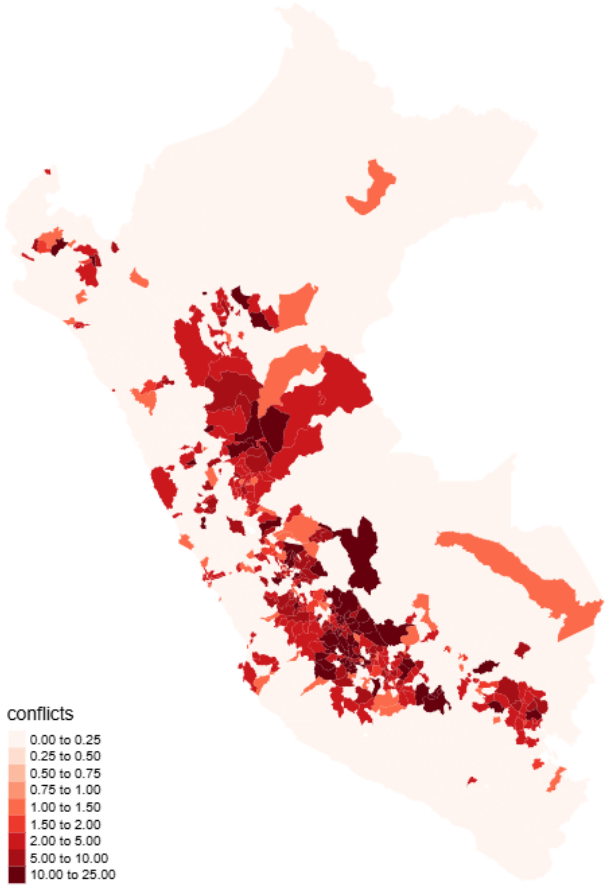


Figure 1. Violent events: 1980-2000

Notes: This figure shows the average number of violent events between 1980 to 2000 at the district level in Peru. Violent events includes murders, disappearances, illegal detentions, kidnappings, forced recruitment, torture, rape, and woundings. The data was collected from Peru 's Truth and Reconciliation Commission.

According to the Truth and Reconciliation Commission (TRC), the recruitment of guerrilla fighters was through coercion, violence, and deception (CVR, 2003). Each community or family had a quota of children that had to be given to the fight, and families that opposed to give their children were killed to later enroll their kids in the guerrilla. The disruption of life of the peasantry and the constant fear of being kidnapped or killed caused discomfort in the rural population, who organized in peasant patrols to fight the Shining Path, which led to an increase of violence after 1983.

Even though the Shining Path conflict was one of the bloodiest in Latin America, the Shining Path considered themselves an "armed group without arms". Their leaders argue that their best arm was their ideology and that the organization did not depend on arms. The Shining Path leaders expected each militant to acquire arms by their own means. Hence, most were acquired by assaulting defenceless policemen or by stealing dynamite from mines.

The conflict extended until 1992 when state forces captured Abimael Guzman. By 1992, the conflict had extended to one-third of all Peruvian municipalities, covering 75% of all departments. Figure 1 shows the spatial distribution of conflict events at the district level in Peru.[3] The figure shows that violence was spread through Peru but was mostly prevalent in the Andean rural region. After the capture of Guzman, violence declined and by 1995 had stopped in most of the regions.

2.2. The Shining Path and Indigenous people

To understand the role of Indigenous people and ethnic identity during the Shining Path insurgency, it is important to understand the Shining Path ideology developed by its leader Abimael Guzman. This ideology was called "Pensamiento Gonzalo", and was an application of communist ideas from Peruvian socialists, like Mairátegui and Maoist ideas into the Peruvian context (Jima-González and Paradela-López, 2020, 2021).

The foundation documents of the Shining Path state that the Shining Path is a "Marxist-Leninist-Maoist" political organization (Shining Path, 1988). Mostly aligned with Mao's ideology, which is relevant to the Peruvian context. According to Mao's ideology, the revolutionary subject is the peasant in contrast to Marxist-Leninist theory, who sees the proletariat as the revolutionary subject. Furthermore, Mao's theory identifies rural areas as the starting point of the revolution and not industrial cities as Marx suggested. Last, due to the anti-revisionism paranoia, high levels of violence and repression were legitimized, specially to those that rejected joining the movement (Jima-González and Paradela-López, 2020).

The Shining Path leaders saw the Indigenous people as the "soldiers of the revolution". According to the TRC, the Shining Path leadership exhibited a sense of superiority toward the Indigenous people (CVR, 2003). For example, in an article by the Peruvian Communist Party [Shining Path] in 1976 they stated: *"We need to build the Party [and revolution] in the countryside, we need to mobilise, organise, and arm the peasants [...] under the direction of the [Communist] Party."* Furthermore, they added: *"The war belongs to the peasants under the leader-*

[3]Conflict events include: killings, sexual assaults, kidnapping and torture.

ship of the proletariat [The Communist Party]. The peasantry, because it constitutes the base of the national democratic revolution, is also the base of the people's war. This is a peasant war, or it is nothing." (PCP, 1976). The Shining Path saw the Indigenous, who compose the majority of the rural population, as soldiers that would allow them to reach their political objectives.

The Shining Path understood that the peasantry and not the Indigenous were the revolutionary subject. Hence, Guzman's ideology, *Pensamiento Gonzalo* included the re-education of the Indigenous people because they were not aware of their own alienation (Jima-González and Paradela-López, 2020). Therefore, the Shining Path emphasised the class identity (peasantry) over any other feeling or identity, such as ethnic or gender identity, because these types of identities were perceived as alienation from capitalism. Along these lines, the Shining Path enforced the abandonment of ethnic [Indigenous] identity to embrace the peasant identity (Jima-González and Paradela-López, 2020).

The insurgency never included any Indigenous values in their official documents. Further, Indigenous people were never part of Shining Path's leadership, which was headed by Guzman and a group of academics. According to the TRC, the Shining Path never took into consideration the cultural differences and identities of the peasantry but considered the peasantry as a homogenous subject (CVR, 2003, p 129).

The use of violence was central in *Pensamiento Gonzalo*, first as a tool for convincing people to join the revolution and second as a way to repress people that did not join them (Jima-González and Paradela-López, 2020). For Guzman, maintaining the "puritiy of the revolution" justifies violence and repression. In a speech to teachers in the Department of Huamanga in 1974, Guzman expressed that: *"war is necessary for the realization of the revolution"* (Gúzman, 1974).

In this context, the Indigenous population had the choice to join the movement and abandon their Indigenous identity and practices to embrace the peasant-class identity or suffer the potentially violent consequences. Anecdotal evidence suggests that this idea was impregnated in the mid-level leaders in the Shining Path. For example, Wheat (1990) pointed out that a leader of the Shining Path organization expressed: *"First they [rural population] will be given political reeducation, and if that fails, they will receive what we get now: dictatorship, prison ... death. The genocide which we have accepted, they will have to accept"* (Wheat, 1990 p. 48).

Following these ideas, The Shining Path started an execution campain against individuals and organizations that refused to collaborate with them (Kent, 1993). Since their main areas of action were the rural areas of Peru, the Indigenous people was the most affected population. Jima-González and Paradela-López (2020) provides anecdotal

evidence of part of the Indigenous communities affected by the Shining Path guerrilla such as the assassination in 1982 of Alejandro Huamán, leader of the Uchuraccay, the killing of 55 Ashaninka peasants in 1993, and the killing of 80 peasants in Lucanamarca in 1983 because they were presumably collaborating with the Peruvian government. In words of Jima-González and Paradela-López (2020), these killings were: *"Justified by Pensamiento Gonzalo's idea that the "blood-quota" was necessary for the triumph of the revolution and warned of the need to prepare for the inevitable bloodbath"* (Jima-González and Paradela-López, 2020, p. 204).

2.3. Identity Formation

There is an extensive literature discussing the consequences of [ethnic] identity (Akerlof and Kranton, 2010). However, a less common question in the economic literature is how identity is formed (Casey and Dustmann, 2010).

The concept of ethnicity and identity has been discussed in depth in the psychological literature. Tajfel and Turner (2004, p. 17) defines [social] *"identity"* as those aspects of an individual self-image that are derived from social categories (e.g. ethnicity, gender, class...) to which she perceives belonging herself. Holcomb-McCoy (2005) define *"ethnicity"* as a set of behaviours, languages, and worldview that are associated with cultural heritage. Hence, ethnic identity is a membership to a specific ethnic group. Furthermore, the psychology literature argues that (ethnic) identity is formed in three different stages (Erikson, 1968; Marcia et al., 1980; Phinney, 1989). First, is a phase where individuals do not explore or do not have a conscious of identity. Next comes a phase of exploration, where individuals actively question and choose among different alternatives. Last, individuals commit to an identity (Phinney, 1989, p. 38).

The (psychologyc) empirical literature suggests that ethnic identity is formed during childhood (0-9) and adolescence (10-19). Brand et al. (1974) shows that ethnic awareness begins at the age of 4. Further, once attitudes towards identity are formed, they tend to increase with age. Davey and Mullin (1980) measure the ethnic identification and ethnic preferences of children aged 7-11. The authors find that children perceive the advantages of being white, and only few children are ethnocentric in the decision to make friends. Marks et al. (2007) study ethnic identification among children of different ages, from 6 to 12 years old. The authors find that older cohorts show greater ethnic identity exploration than younger cohorts. The majority of studies focus on the development of ethnic development during adolescence.[4]

[4]for a review, see: Phinney (1990)

Seminal work by Phinney (1989) show that about half of the adolescent have not explored their ethnic identity, one quarter is exploring their identity and one quarter have achieved their identity. Yip (2014) finds that adolescents that achieved ethnic identity have a high level of identity salience across different situations.

Following the literature on the identity formation process, we hypothesize that if exposure to violence affects ethnic identity, it would affect individuals exposed to violence between the ages of 0 and 19, when identity is formed. We expect that exposure to violence do not affect ethnic identity formation of individuals exposed to violence after the age of 20.

The direction of the consequences of violence on ethnic identity is unclear. On the one hand, exposure to violence could decrease generalized trust, leading individuals to shelter in their ethnic group and reinforcing ethnic identity. This is relevant for inter-ethnic group violence (Rohner et al., 2013a,b). On the other hand, we hypothesize that exposure to intra-ethnic group violence could hamper individuals from self-identifying as their ethnic group following an opposite dynamic as suggested by Rohner et al. (2013a,b).

3. Empirical framework

3.1. Data

Dependent Variable. The dependent variable is an individual's ethnic identity, more specifically whether an individual is Indigenous or not. We measure ethnic identity in two ways. First, we measure Indigenous identity based on individuals' self-identification whether an individual self-identify as Indigenous or not, and second based on their mother tongue.

We collect data at the individual level from the *Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza* (ENAHO) originally collected by *Instituto nacional de Estadística e Informática* (INEI). This survey collects individual-level data on the living conditions of people across all 24 Peruvian departments (first administrative unit). We use the survey waves from 2013 to 2017 because these waves include questions regarding individuals' Indigenous identification and mother tongue. The surveys are organized in 29 different modules where the INEI collects information regarding the livelihood of Peruvians. For the analysis, we use the following modules: Module 1: house characteristics, (*Características de la Vivienda y del Hogar*); Module 2: charateristics of individuals (*Características de los Miembros del Hogar*); Module 3: Education (*Educación*); Module 4: Health (*Salud*); Module 5: Income and employment (*Empleo e Ingresos*); ands Module 85: governance, democracy, and transparency (*Gobernabilidad, Democracia y*

Transparencia).

We collect the dependent variable from the education, and employment and income modules. In the employment & income module, we can find the question: *"According to your ancestors and your customs, do you consider yourself: Quechua, Aymara, Other Indigenous community, Black, White, Mestizo, or unknown?"* We classify an individual as Indigenous if the individual answered Quechua, Aymara and Other Indigenous community. We classify an individual as non-Indigenous if the Individual answer mestizo. The main reason we only use mestizo as a base is historical. Mestizo is a person with both Spanish and Indigenous ancestry, hence, if an Individual changes their ethnic identity, it would move from Indigenous to mestizo or mestizo to Indigenous rather than to white or black.

From the education module, we use the question: *"What is the language or mother tongue you learned in your childhood?"* The potential answers were: Quechua, Aymara, another Indigenous language, Spanish (Castellano), English, Portuguese, another foreign language, sign language or other. In this case, we classify an individual as Indigenous if the individual answered Quechua, Aymara, or another Indigenous language. If the individual answered Castellano, we would classify them as non-Indigenous. Similarly, as the previous case, Castellano (Spanish) is the most common language in Peru, hence, if an individual did not learn an Indigenous language, Castellano would be the most likely language to learn.

Figure A.1 in the appendix shows the share of individuals each cohort from 1950 to 1992 by ethnicity. In subfigure A.1a, we use mother tongue as a measure of ethnicity and in subfigure A.1b, we use the self-identification measure. The figure shows an increase of Non-Indigenous identification in younger cohorts. This could be a result of the violent consequences of the Shining Path conflict.

Independent Variable. We collect data on Peru's conflict from the Truth and Reconciliation Commission (TRC). This dataset provides data on violent events at the event level. It provides information about the location, time, type of violence and victim of the event from the beginning of the conflict in 1980 until 2000.

The TRC collected 36,019 unique reports of violent events from testimonies of human rights abuses through interviews and visits to communities affected by the conflict. The conflict events came from 16,917 testimonies and included murders, disappearances, illegal detentions, kidnappings, forced recruitment, torture, rape, and woundings.

To construct the independent variable, we combine the conflict event dataset from TRC and the individual-level dataset from ENAHO. Using the location and time of the event, we construct a panel dataset that counts the number of conflict events in each of the 1873 Peruvian districts from 1980 until 1992. Further, the ENAHO dataset provides in-

formation on individuals' district and year of birth (Module 2). We combine these two datasets and create the variable *exposure*, which takes the value of one if an individual was born in a district that suffered a violent event and takes the value of zero otherwise. Furthermore, the dataset allows us to differentiate whether individuals were exposed to violence in their district of birth during their ethnic identity formation years (0-19) or when ethnic identity was already formed (>19). Last, since we are interested in assessing the impact of violence on Indigenous identity, we consider only violent events with an Indigenous person as the victim. According to the TRC, approximately 75% of the victims of the conflict spoke Quechua or another Indigenous language as their first language (Laplante and Theidon, 2007).[5] Table A.1 in the appendix presents the summary statistics of all the variables used in the empirical analysis.

3.2. Sampling design

The main concern of the empirical strategy is that the Shining Path did not randomly select districts to perpetuate their violent attacks. Figure 1 shows the spatial distribution of conflict events in Peruvian districts from 1980 to 2000. The figure shows that most Shining Path attacks were concentrated in the Andean region of Peru and the Shining path was mostly present in rural areas. Even though Indigenous people were not the main target of the Shining Path, the Indigenous population were the main victims, and they represented approximately 75% of the total death of the conflict (Laplante and Theidon, 2007).

To investigate the effect of the Shining Path conflict on ethnic identity, the ideal experiment would be to randomly allocate Shining Path's terrorist attacks in Peruvian districts and compare whether individuals who were born before and after the attacks in districts with and without attacks self-identify ethnically different. However, this field experiment is not feasible. Hence, to provide evidence for the consequences of conflict on ethnic identity, we select a set of bordering districts that were, arguably, similar before the conflict started but suffered different violence intensity. A similar strategy was used by Lichter et al. (2021), where the authors use a border discontinuity design to provide evidence for the long-run effects of government surveillance on civic capital and economic performance, using East Germany district borders.

The objective of this section is to find a set of districts that were equal in socio-economic characteristics before the Shining Path conflict started but suffered different levels of violence due to a random

[5]Note that the variable could suffer from measurement error, if an individual was born in a district that suffered a conflict event but the individual moved to a different district (false positive) or if an individual was born in a district that did not suffer a conflict but moved to a district that expose the individual to a conflict event (false negative).

characteristic. The sampling design is based on Albertus (2020) study. The author studies the effect of land reform on conflict in Peru. Using a regression discontinuity design, the author finds that districts in core areas of land reform witnessed less conflict than the adjacent districts in peripheral areas where less land reform occurred.

Figure 2 shows the standardized mean differences coefficients of different variables between core and periphery districts. The mean comparison indicates no differences between core and periphery districts in the vote for Marxist political parties (coef 3), and there are no differences in the share of indigenous people (coef 4) before the conflict started. Further, the results show that core municipalities suffer a higher number of conflict attacks and a higher number of deaths in the conflict, coef (1) and (2), respectively. These results suggest that when using the bordering discontinuity sample, conflict events and deaths were not confounded by the voting to Marxist or the share of Indigenous people. Last, we do not find statistical differences in geographical and socio-economic characteristics, such as population, the literacy rate in 1961 and 1972, the number of social movements prior to the conflict, the number of hectares of private land, the share of cultivated land in each district, the number of state employees and land reform. We find that core districts had a lower elevation and slope than periphery districts. The latter can be controlled with district-fixed effects.[6] Given these results, we use the bordering sample as the main sample in the anaylsis.

[6]In appendix D see a discussion about the differences between the sample design in this paper and the RDD in Albertus (2020).

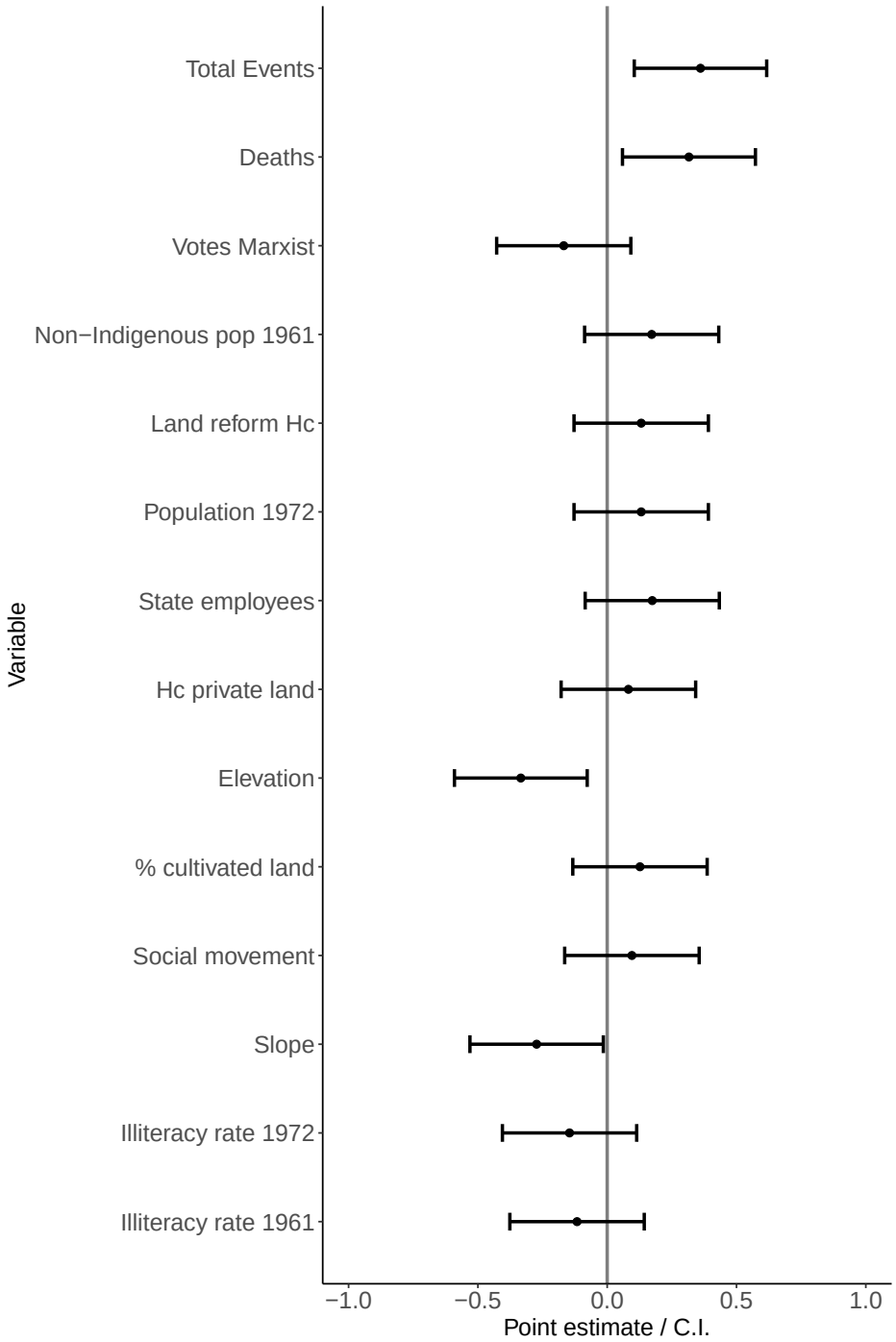


Figure 2. Mean comparison

Notes: The figures plots the standardized mean differences of different variables between core and periphery districts with 95% confidence intervals. The base category is core districts. The unit of observation is the districts. The variables are total number of violent events, total number of death, share of votes to Marxist parties , share of non-Indigenous population in 1961, illiteracy rate in 1961, illiteracy rate in 1972, peasant movements, average slope, share of cultivated land, Hectares of private land, average elevation, number of state employees, population in 1972 (thousands), Hectares of land reform. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

3.3. Empirical strategy

The main objective of this paper is to study the effect of violence on ethnic identity. To capture the impact of the Shining Path violence on ethnic identification, we assume that the district and year of birth determine the exposure to violence of individuals. Furthermore, we use the border sample to select a set of districts which are similar in different socio-economic characteristics, except for the level of violence intensity, to provide quasi-random variation in the exposure to violence of individuals. Last, we compare ethnic identification differences between individuals who were exposed and not exposed to violence when they were between 0 and 19 years old with respect to the ethnic identification differences of individuals who were exposed and not exposed to violence when they were older than 19 years old. We implement this empirical strategy in a difference-in-differences regression framework applying the following specification:

$$Y_{ikdtp} = \beta Exposure_{ip} \times T_{ipk} + \gamma_k + \gamma_t + \gamma_d + \gamma_p + \delta c_i + \varepsilon_{iktp} \quad (1)$$

Where Y_{ikdtp} is a dummy that captures the ethnic identity $\in$(Indigenous or non-Indigenous) of individual i, born in year k, living in district d, interviewed in year t, and born in district p. We use two measures of ethnic identity, *Identify Indigenous*, which takes the value of 1 if the individual identifies as Indigenous and 0 if the individual identity as mestizo. *Mother tongue*, which takes the value of 1 if an individual speaks an indigenous language as his mother tongue and 0 if the individual speaks Castellano (Spanish) as his mother tongue.

$Exposure_i$ takes the value of 1 for individuals that were exposed to violence in their district of birth and 0 otherwise. T_{idk} takes the value of 1 for individuals exposed to violence between the age of 0 and 19 and 0 otherwise. $\hat{\beta}$ is our coefficient of interest, which captures the difference in ethnic identification between individuals who were exposed and not exposed to violence when they were between 0 and 19 years old with respect to the ethnic identification differences of individuals who were exposed and not exposed to violence when they were older than 19 years old. γ_k are year-of-birth fixed effects. γ_t are year-at-interview fixed effects, γ_d are district-at-interview fixed effects, and γ_p are district-of-birth fixed effects. δc_i are individual-level controls, such as age, age^2, and sex.

This econometric setting allows us to mitigate potential endogeneity concerns. Year-of-birth fixed effects allow us to compare individuals within the same year of birth and control for a possible decrease in Indigenous identification over time. District fixed effects control for time-invariant factors such as geography. Year-at-interview fixed effects allow us to control for shocks that occurred during the year of

the interview, for example, nationwide economic recession. District-of-birth FE allows us to compare individuals that were born within the same district. Furthermore, our border sample allow us to select districts that were similar in socio-economic characteristics, like the share of Indigenous population, and political views, but suffered different levels of violence due to an exogenous border design. Our main identifying assumption is that, in the absence of violence, cross-cohort ethnic self-identification would have been similar between cohorts with different levels of violence exposure.

We use a more flexible difference-in-differences strategy to understand the consequences of violence on ethnic identity. In this specification, we compare how exposure to violence at different age cohorts impacts ethnic identity. Furthermore, this specification allows us to compare individuals' ethnic identification born after violence was perpetrated in their district of birth compare to the baseline cohort (20-24). To do so, we estimate the following regression equation:

$$Y_{ikdtp} = \sum_{c} \beta_c Exposure_{ip} \times D_{pc} + \gamma_k + \gamma_t + \gamma_d + \gamma_p + \delta c_i + \varepsilon_{ikdtp} \tag{2}$$

Here, D_c is an indicator whether individual i experienced violence when belonging to cohort $c \in$ (post, 0-4, 5-9, 10-14, 15-19, 25-29, 30-34). Our baseline cohort is those individuals that were between 20 and 24 when violence occurred in municipality p. The rest of the variables remain the same as in our previous specification. The coefficients of interest, $\hat{\beta}_c$, capture cohort-specific treatment effects. More specifically, they capture the ethnic identification differences between individuals within each cohort c that were exposed and not exposed to violence with respect to the baseline.

4. Results

Table 1 shows the results of our specification embedded in equation 1. The dependent variable in columns (1)-(2) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (3)-(4), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (3), we use the full sample of individuals. In columns (2) and (4), we restrict the sample to the districts in the border sample. In all specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Additionally, we control for individual characteristics, such as age, age square, and sex. Last, we cluster the standard errors at the district-year level.

The results show a negative association between exposure to violence and individuals speaking an indigenous language as their mother

tongue, as well as identifying as Indigenous. When using the full sample, individuals that were exposed to violence are 1.4 percentage points (pp. henceforth) less likely to speak an indigenous language than similar individuals that were exposed to violence during the identity formation years. Further, there is no statistical significant association between exposure to violence and self-identifying as Indigenous.

When using the border sample, the negative association between exposure to violence and Indigenous identity become larger in absolute terms. The results show that individuals exposed to violence are 3.4 pp. less likely to speak an indigenous language than individuals that did not experienced violence during their ethnic identity formation years. Similarly, individuals exposed to violence are 2.3 pp. less likely to identify Indigenous compared to individuals that were not exposed to violence. The effect is non-negligible in size, approximately 25% of individuals in the border sample speak an indigenous language as their mother tongue and 40% identify Indigenous. Hence, a 3.4 pp. (2.3 pp.) decrease in the likelihood of speaking an indigenous language (identifying as Indigenous) represents a 13.6% (5.75%) with respect to the mean (40%).

Table 1. Total

	Indigenous language as mother tongue		Identify Indigenous	
	all districts	border sample	all districts	border sample
	(1)	(2)	(3)	(4)
Exposure violence x T	−0.014**	−0.034**	−0.005	−0.023**
	(0.006)	(0.016)	(0.005)	(0.009)
District-interview FE	✓	✓	✓	✓
District-born FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DP	0.23	0.25	0.36	0.40
SE Cluster	district-year	district-year	district-year	district-year
N	208,768	33,326	168,301	27,842
Adjusted R^2	0.660	0.628	0.545	0.480

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1)-(2) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (3)-(4), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (3), we use the full sample of individuals. In columns (2) and (4), we restrict the sample to the districts in the border sample. In all specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

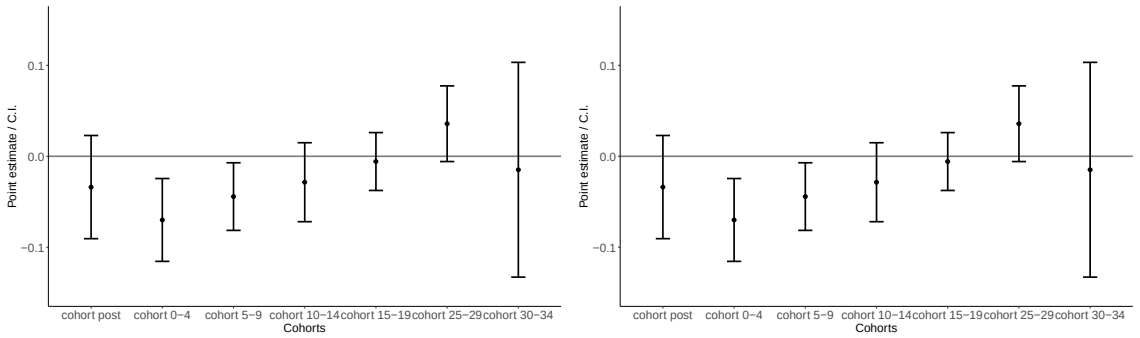
These results are robust to different specifications. Tables A.3 and A.2 in the appendix are replicas of table 1, but in table A.3 we remove the cohort that never was exposed to violence, and in table A.2 we cluster the standard errors at the district of birth.

We perform cohort analysis to understand which cohorts were more affected by violence. Figure 3 plots the β_c coefficients of equation 2. This figure shows differences in ethnic identification between individuals that were exposed to violence at different ages with respect to differences in ethnic identification of our baseline cohort (20-24 years old). Furthermore, this specification allows us to provide evidence regarding the assumption that Indigenous identification would be similar between cohorts in the absence of violence. To do so, we compare the ethnic identity of individuals exposed to violence, after the identity formation years, between the ages of 25-29 and 30-35 with our baseline (cohort 20-24). Further, we compare the ethnic identity of individuals born after a violent event occurred in a district (*post* cohort) with respect to the baseline cohort.[7] Table A.6 in the appendix shows the full set of results of this analysis.

The results show that ethnic identity is mostly affected when individuals are exposed to violence during childhood (0-9 years old). The estimates suggest that individuals exposed to violence between 0 and 4 years old are 7 pp. less likely to speak an Indigenous language as their mother tongue and 4 pp. less likely to identify Indigenous than individuals that were not exposed to violence. Furthermore, individuals that were exposed to violence between the age of 5 and 9 are 4.4 pp. less likely to speak an Indigenous language as their mother tongue, but the likelihood of identifying as Indigenous is not different from zero. Last, the results show no statistical differences in the probability of speaking an Indigenous language as a mother tongue or identifying as Indigenous among individuals that were exposed to violence between the age 10 and 19.

To provide evidence for the assumption that ethnic identification is not different between cohorts in the absence of violence, we compare ethnic identification differences between cohorts exposed to violence when their identity was already formed (cohorts 25-29, and 30-34) and the baseline cohort. The results show no statistical differences in Indigenous identification between cohorts 25-29, and 30-34, with respect to our baseline. Last, we find no differences in Indigenous identification between individuals born after violence was perpetrated (*post* cohort) in the district they were born and our baseline cohort.

[7] Post cohort example: an individual born in the year 1987 in district i when the last violent event reported in district i was in 1984.



(a) Mother tongue (b) Indigenous identification

Figure 3. Cohort-level analysis

Notes: The figure plots β_τ coefficients estimated from equation 2 with 95% confidence intervals. The omitted cohort is 20-24. The explanatory variable is the interaction term between exposure to violence and exposure to violence after 19 years old. In subfigure 3a the dependent variable is whether an individual speaks an Indigenous language as their mother tongue or not. In subfigure 3b is the dependent variable is whether an individual identifies Indigenous or not. We restrict the sample to the districts in the border sample. In both specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level.

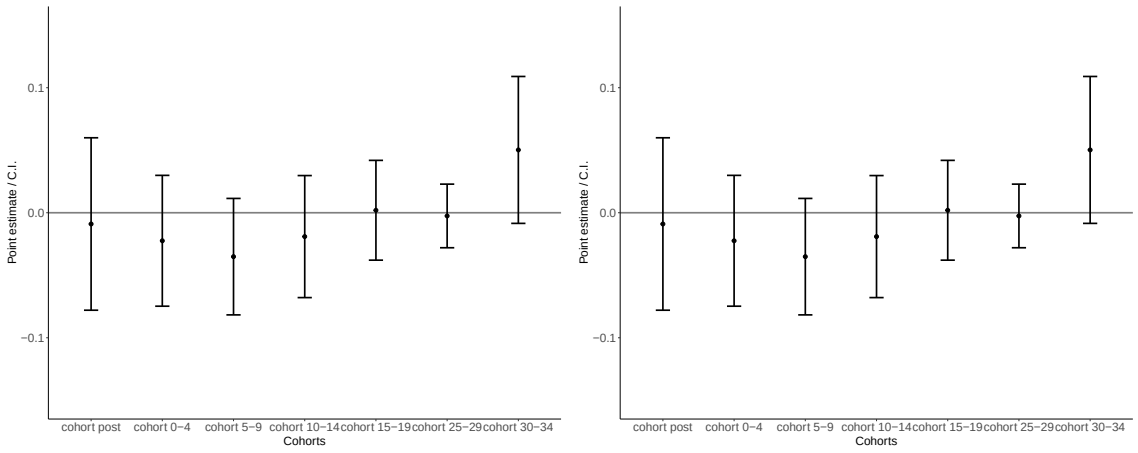
These results suggest that the Shining Path violence affected ethnic identity in Peru. However, the conflict did not affect all individuals equally. Individuals who were exposed to violence during childhood were less likely to speak an indigenous language as their mother tongue and less likely to identify as indigenous.

The attachment theory from the psychological literature could help understand why violence affected the development of an Indigenous identity during infancy and early childhood but not adolescence (Ainsworth, 1989; Bowlby, 1982). This theory argues that individuals during childhood and later in life have attachment figures for their emotional and social development, which includes the identity formation process (Parkes et al., 2006). During infancy and childhood, the main attachment figure is the mother. However, during adolescence, identity formation is not only influenced by parents but also by others in society (e.g. teachers in school) (Cama and Sehgal, 2021). The Shining Path ideology forced individuals to embrace a class identity abandoning all other identities, including ethnic identity. This could have led parents to stop transmitting their languages and culture to children fearing potential violent consequences. When violence occurs during adolescence, children could have already learned an Indigenous language during their childhood, mitigating the negative consequences of violence on Indigenous identity.

4.1. Robustness checks

4.1.1. Non-Indigenous violence.

This paper shows that individuals exposed to violence are less likely to identify Indigenous and to speak an Indigenous language as their mother tongue. In the main analysis, we use only violent events where the victim was an Indigenous person (approximately 77.3% of the violent events).



(a) Mother tongue (b) Indigenous identification

Figure 4. Non-Indigenous violence

Notes: The figure plots β_τ coefficients estimated from equation 2 with 95% confidence intervals. The omitted cohort is 20-24. The explanatory variable is the interaction term between exposure to non-Indigenous violence and exposure to non-Indigenous violence after 19 years old. In subfigure 4a the dependent variable is is whether an individual speaks an Indigenous language as their mother tongue or not. In subfigure 4b is the dependent variable is whether an individual identifies Indigenous or not. We restrict the sample to the districts in the border sample. In both specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level.

In this section, we use violent events where the victim was not an Indigenous person and test whether exposure to a violent event of a non-Indigenous person affects Indigenous identity. To do so, we use the cohort analysis and compare Indigenous identification differences between Individuals that were exposed and not exposed to non-Indigenous violence with respect to the baseline difference (20-24 cohort).

Figure 4 shows the coefficient of interest (β_c) of this event study analysis. Subfigure 4a shows that exposure to non-Indigenous violence has no effect on the probability of an individual speaking an Indigenous language as a mother tongue. Subfigure 4b shows that exposure to non-indigenous violence has no consequences on the likelihood of an individual identifying as Indigenous. These results suggest that violence mainly affects Indigenous identification when the victim is an

indigenous person.

4.1.2. Migration

The economic literature has shown a relationship between conflict and migration.[8] In the full sample, the data shows that individuals that were exposed to violence are 8.4 pp. more likely to migrate than individuals that were not exposed to violence (see column in 7 in table A.4 in the appendix). Hence, migration could be a potential confounder. The datset set allow us to control for the potential biases of migration.

In the main specification, individuals are exposed to violence if there was a violent event in their district of birth when they were between 0 and 19 years old. We do not differentiate between individuals that migrated and individuals that stayed in their district after a violent event. To further understand whether individuals that migrated and did not migrate have different ethnic identification after a violent event, we create a dummy variable, *migration*, which takes value 1 if the individual is living in a different district, at the time of the interview, than the one he was born and 0 if he lives in the same district as he was born. Furthermore, we use the main identification strategy embedded in equation 1 and include the *migration* in different ways.

Table A.4 in the appendix shows the results when including the *migration* dummy. In columns (1)-(3), the dependent variable is whether an individual speaks an Indigenous language as mother tongue. In columns (4)-(6), the dependent variable is whether an individual identifies Indigenous or not. Columns (1) and (4) are our baseline results similar to columns (2) and (4) in table 1. In columns (2) and (5), we include the migration dummy as a control variable. In columns (3) and (6), we interact the migration dummy with our measure of exposure to violence. In all regressions, we include district fixed effects, year-interview fixed effects, and year of birth fixed effects. Further, we include the individual controls, sex, age, and age square. Last, the standard errors are clustered at the district-year level.

The results show that when controlling for migration, the results are not statistically different from our baseline. The results show that individuals that were exposed to violence are 3.4 pp. less likely to speak and Indigenous and 2.3 pp. less likely to identify as indigenous than comparable individuals not exposed to violence. The interaction term results show that violence mostly affected individuals that were exposed to violence and did not migrate. The baseline comparison group is no-exposed to violence and no-migration, hence, each coefficient is compared to this group. The results show that individuals that were

[8]See: Guatemala (Morrison, 1993); Mexico (Basu and Pearlman, 2017); Colombia (Silva and Massey, 2015); Africa (De Vries and Guild, 2019)

exposed to violence and did not migrate are 5.1 pp. less likely to speak an Indigenous language as their mother tongue and 3.6 pp. less likely to identify as Indigenous than Individuals that were not exposed to violence and did not migrate. Furthermore, the results show no statistical differences in ethnic identification between individuals that were not exposed to violence and migrated and individuals that were not exposed to violence and did not migrate. Last, we do not find a difference in the likelihood of speaking an indigenous language as a mother tongue or indigenous self-identification between individuals exposed to conflict and migrated and individuals not exposed to violence and not migrated.

To further understand the consequences of the Shining Path violence on migration, we use the *migration* dummy as an outcome using equation 1. Column (8) in table A.4 shows that the probability of migrating is not different for individuals exposed to violence before and after age 20. These results show that the migration was not a confounder driving the the main results.

4.1.3. Other identities

This paper studies the consequences of violence on ethnic identity. As explained in section 2, the Shining path started its revolutionary activity and pressured the population to embrace a class identity. In this section, we study the consequences of violence on other types of identities, namely class identity and national identity.

To test the consequences of violence on class identity, we collect data on individuals' membership in different class-related organizations from the ENAHO surveys. To the best of our knowledge, there is no survey data on self-identification on class identity in Peru. Therefore, we use membership in workers and professionals associations, membership in political parties, and membership in a peasant community as a proxy for class identity. We expect that if class identity is salient in an individual, the likelihood of being part of one of these organizations is higher compared to an individual that does not emphasize class identity.

We use the specification in equation 1 and replace the dependent variable with a dummy variable indicating the probability of an individual belonging to a worker association, a professional association, a political party organization, and a peasant community. Table 2 show the results of this analysis. The estimates suggest that exposure to violence has no consequence on the probability of an individual participating in a workers' or professionals' association nor in the likelihood of being part of a political party organization, which suggest that the exposure to the Shining path violence had no consequences on class identity. Further, we find that exposure to violence decreases an indi-

vidual's probability of being part of a peasant community. Note that the Shining Path ideology pressured the individuals to embrace a peasant identity. These results suggest that the exposure to their violence had the opposite effects. The likelihood of belonging to a peasant association is lower when an individual is exposed to the Shining Path's violent activity.

To test for the consequences of violence on National identity, we collect data on national identity from the Latinbarometro survey.[9] From this database, we use the question: *"How proud are you to be Peruvian?"* The respondents could answer: 1) very proud, 2) proud, 3) little proud, and 4) not proud. From this question, we create a dummy variable which takes the value of one if the respondent answered very proud or proud and the value of 0 if the respondent answered little proud or not proud. Further, we create a ordered variable, which takes values of 1, 2, 3 and 4 if the respondent answered not proud, little proud, proud, and very proud, respectively.

We combine the Latinbarometro survey and the TRC database and use the specification in equation 1 to understand the consequences of exposure to violence on National identity. Table 3 shows that individuals exposed to violence are less likely to be proud of their Peruvian identity than individuals that were no exposed to violence (columns 1 and 3). The results remain similar when controlling for indigenous identity (Columns 2 and 4). These results suggest that the Shining Path violence had a consequence on national identity, which could be attributed to the lack of state intervention, especially in the early years of the conflict.

4.1.4. Ethnic group

Peru is a multiethnic society. In the main specification, we only compare Indigenous and mestizos. However, in Peru, there are more than 50 different Indigenous groups. In our dataset, we have Quechua, which represents 22% of the individuals in our sample, Aymara (0.6%) and other Indigenous populations (2.9%). In this section, we check whether exposure to violence has heterogeneous effects depending on the affected Indigenous group. To do so, we use the main specification but splitting the sample into Quechua, Aymara, and other Indigenous groups. Note that our baseline sample is always Castellano for the mother tongue language variable and mestizo for the self-identification variable.

Table A.5 presents heterogeneous effects of the association between exposure to violence and different Indigenous identity. The results show that individuals exposed to violence were less likely to speak

[9] In section 5.1, we discuss the potential shortcomings of using the Latinbarometro survey.

Table 2. Class-identity

	Workers	Professionals	Political Party	Peasant community
	(1)	(2)	(3)	(4)
Exposure violence x T	0.008	0.005	0.002	−0.023**
	(0.008)	(0.009)	(0.005)	(0.011)
District-interview FE	✓	✓	✓	✓
District-born FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DV	0.037	0.37	0.011	0.22
N	13,868	13,868	13,868	13,868
Adjusted R^2	0.084	0.062	0.002	0.641

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1),(2), (3), (4), are whether an individual belongs or not to a workers association, a professionals associations, a political party, and a peasant community, respectively. In all specifications, we include, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table 3. National identity

	National identity dummy		National identity	
	(1)	(2)	(3)	(4)
Exposure violence x T	−0.057***	−0.052**	−0.099***	−0.098**
	(0.021)	(0.020)	(0.036)	(0.038)
No-Indigenous		0.084***		0.177***
		(0.027)		(0.044)
District FE	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DV	0.83	0.83	3.38	3.38
SE Cluster	district-year	district-year	district-year	district-year
N	2,675	2,656	2,675	2,656
Adjusted R^2	0.010	0.010	0.024	0.024

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1)-(2) is the ordered measure of national identity. In columns (3)-(4), the dependent variable is the dummy variable measure of national identity. In columns (2) and (4), we include a dummy variable indicating whether the individual is Indigenous or not. In all specifications, we include, year-birth fixed effects, year-survey fixed effects, district fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Quechua or Other Indigenous languages as a mother tongue, and less likely to identify as Quechua and Other indigenous groups than comparable individuals that were not exposed to violence in the identity formation years compared to differences between individuals exposed and not exposed to violence after the identitiy formation years. The results regarding the Aymara population are not statistically different from zero, due to the lack of statistical power. Hence, it is not possible to draw a conclusion on the consequence of violence with respect to this Indigenous group.

The majority of the Indigenous population in Peru is Quechua, representing around 22% of the total population, whereas the Aymara population represents less than 1%. We hypothesize that the negative association between exposure to violence for the Quechua population but not the Aymara population is own-Indigenous group violence. Since the Quechua population is the majority, the likelihood of experiencing own-group violence is higher than for the Aymara population, leading to a lower Indigenous identification. This mechanism is explained in the next section.

5. Dynamics of Conflict and Mechanism

The results show that those individuals exposed to the shining path violence in their identity formation years have a lower probability to identify indigenous and a lower probability of speaking an Indigenous language as a mother tongue. These results are particularly interesting in light of recent advances in understanding the dynamics of conflict and ethnic identity.

Rohner et al. (2013b) propose a theoretical model to understand the dynamics of violent conflict, trust and ethnic identity. The authors suggest that civil wars are persistent over time, especially in multi-ethnic societies. The theory suggests that conflict and violence erodes trust in a society, leading individuals to identify more with their own ethnic group and increasing the probability of future conflicts, plunging the society into a vicious cycle, which the authors describe as a *"war trap"*. To provide evidence for their theoretical predictions, Rohner et al. (2013a) use the civil conflict in Uganda between 2002 and 2005. The findings show that in regions where fighting was more intense, individuals decreased generalize trust and increased ethnic identity and interpersonal trust.

In Peru, we do not see a persistent *"war trap"* or persistent violent events towards identity as suggested by Rohner et al. (2013b). Figure 5 shows that the Peruvian conflict was short and not persistent. It started to spike in 1983 until 1985. Between 1986 and 1987 there was a period of low conflict intensity, and it rose again from 1988 un-

til Abimael Guzman was captured in 1992. After 1995, the conflict was negligible and not persistent. The main difference between the Peruvian civil conflict and the theoretical predictions by Rohner et al. (2013b) are the perpetrators of the violent attacks. In the Peruvian case, the perpetrators were individuals from the same village as the victims, whereas in Rohner et al. (2013b), they were from different ethnic groups. Therefore, we hypothesize that the mechanism driving the negative relationship between exposure to violence and Indigenous identity is own-ethnic group violence.

To understand the dynamics of violent conflict on trust and ethnic identification, we follow Rohner et al. (2013a,b) and study the consequences of the Shining Path violence on generalized and interpersonal trust. Furthermore, we study the consequences of own group violence on ethnic identification.

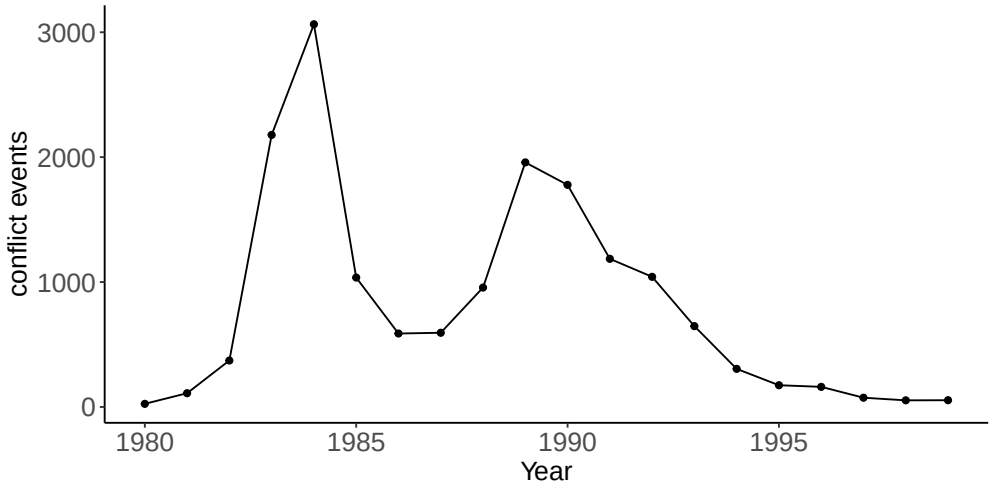


Figure 5. Violent Conflict events by year (1980-2000)

Notes: This figure shows the average number of conflict violent events between 1980 to 2000 in Peru. Violent events includes murders, disappearances, illegal detentions, kidnappings, forced recruitment, torture, rape, and woundings. The data was collected from Peru 's Truth and Reconciliation Commission.

5.1. Trust

Interpersonal Trust. One of the primary mechanisms explaining the dynamics of war is that conflict increases interpersonal trust. Hence, individuals tend to trust their neighbours and relatives more than individuals out of their community. To understand the consequences of the exposure to violence on interpersonal trust, we collect data on interpersonal trust from Latin American Public Opinion Project surveys (LAPOP).[10] We use the LAPOP waves of 2006, 2012, 2014 and 2016. From these surveys, we use the variable *IT1*, which asks the

[10]Source: The AmericasBarometer by the LAPOP Lab, here

question: *"...And speaking of the people from around here, would you say that people in this community are very trustworthy, somewhat trustworthy, not very trustworthy or untrustworthy...?"*. This variable takes value 1 if the individual answered very trustworthy, 2 for somewhat trustworthy, 3 for not very trustworthy, and 4 for untrustworthy. Note that for a more straightforward interpretation of the coefficients, we reverse the original value, hence, the variable takes value 1 for untrustworthy, 2 for not very trustworthy, 3 for somewhat trustworthy, and 4 for very trustworthy.

The use of LAPOP survey comes with different shortcomings, however, to the best of our knowledge, there is no other survey data on interpersonal trust in Peru. The first shortcoming comes from the fact that the data is sampled at the province level instead of the district level. The province level is the second administrative layer in Peru, and the Peruvian territory is divided in 196 provinces. Therefore, when constructing the exposure to violence using this data, individuals were in provinces that were affected by the conflict are classified as exposed to violence independently whether their district recorded a violent event or not. The consequence of this measurement error would push the coefficient of interest towards zero, downward biasing the estimation, since individuals that were not affected by the treatment (violent event exposure) are now "treated".

The second shortcoming is the lack of information about the individuals' district or province of birth. In this case, we make the assumption that an individual province of residence is their province of birth. The main results show that individuals exposed to violence and did not migrate are driving our results. The consequence of being unable to differentiate between migrants and not migrants is that the coefficient of interest could be potentially biased towards zero because those individuals who migrated from provinces exposed to violence are now "treated", and we cannot control for it. These shortcomings of using the LAPOP surveys dataset push our coefficient to zero, hence, in the absence of this classical measurement error, we would expect a higher absolute number of the coefficient with lower standard errors.

To test whether exposure to the shining path conflict affected interpersonal trust, we use the specification in equation 1. In this case, the dependent variable is my measure of interpersonal trust, and the unit of treatment is the province-year due to the lack of data at the district level. Table A.8 in the appendix show the results on interpersonal trust. Column (1) shows that violence affected individuals' interpersonal trust. The result shows that an individual that was exposed to violence had 6.5% (0.167 units) lower interpersonal trust with respect to the mean than individuals that were not exposed to violence.

In column (2), we study whether the effect of trust was different between Indigenous and non-Indigenous individuals. To do so, we create

a dummy (*No-Indigenous*) which takes value 1 for non-Indigenous and zero for Indigenous individuals. Note that the inclusion of this variable could lead to potential bias due to the inclusion of a "bad control", hence, the results should be interpreted with caution as a mean comparison and not as a causal effect. In column (2), we include in the *No-Indigenous* dummy as a regressor. The result shows that conditionally on the ethnic group, individuals exposed to the violence had a lower interpersonal trust than those not exposed to violence. Additionally, Non-Indigenous individuals had a 4.4% (0.11 units) higher interpersonal trust with respect to the mean (2.46) than Indigenous people.

These results show that individuals exposed to violence had lower interpersonal trust than individuals that were not exposed to violence. In contrast to Rohner et al. (2013b), we find that violent conflict decreased ethnic identity and interpersonal trust specially to Indigenous people. This could be the reason for the not persistent wars in Peru after the Shining Path conflict.

Generalized Trust. Rohner et al. (2013a,b) argue that violent conflict erodes generalised trust, leading individuals to identify more with their own ethnic group. To understand whether the Shining Path conflict affected generalized trust, we collect data from the Latinobarometro surveys. We use the wave from 2002 to 2015 and select the question: *"Generally speaking, would you say that you can trust most people, or that you can never be too careful when dealing with others?"*. The respondents could answer: 1) You can trust most people or 2) One can never be too careful when dealing with others. Using this information, we create the variable generalize trust, which takes value 1 if the respondent answered: *"You can trust most people"*, and 0 if the respondent answered: *"One can never be too careful when dealing with others"*.

To study how exposure to violence affects an individual's probability of trusting most people (generalized trust), we use a similar identification strategy to the one studying interpersonal trust. Here, we combine the Latinbarometro survey with the TRC conflict database at the province level and use the specification in equation 1. Note that similarly to the LAPOP surveys, the Latinbarometer surveys are representative at the province level and do not account for migration. Hence, the identification strategy will suffer the same shortcomings.

Table A.9 in the appendix is a replica of table A.8 but the dependent variable is the probability of an individual answering: *"You can trust most people"*. In column (1), we only include the exposure to conflict variable. The results show that individuals exposed to violence are 1.4 pp. more likely to trust someone than individuals that were not exposed to violence. In column (2), we include the *No-Indigenous* dummy as a control. The results are similar and suggest that individuals exposed to

violence are 1.4 pp. more likely to trust someone than individuals that were not exposed to violence. Note that these results are not statistically different from zero.

To further understand how exposure to violence affected trust, we collect data on individual institutional trust from the ENAHO survey module 85 from the waves 2013 to 2017. We collect information about individuals' trust on different institutions. The survey asks the question:*Currently, do you have trust in the following institutions?* The individual could rank from 1 for no trust, 2 for little trust, 3 for enough trust, 4 for a lot of trust. We select trust in the following institutions: district government, national government, the judiciary power, the national congress, and the catholic church. Furthermore, we select a variable that measures whether an individual knows what a democracy or not is.

To test whether the Shining path conflict had an association with institutional trust in Peru, we use the empirical specification embedded in equation 1. We use as a dependent variable our trust measures to different institutions, namely trust to: the district government, the national government, the judiciary power, the congress, and the catholic church. Furthermore, we compare whether individuals exposed to the conflict have a higher probability of knowing what democracy is than individuals not exposed to conflict.

Table A.10 in the appendix show the results of the consequences of exposure to conflict on institutional trust. In column (1), the dependent variable is confidence to the district government. In columns (2), (3), (4), (5), (6), the dependent variables are confidence to the national government, the judiciary power, whether an individual knows what democracy and confidence in the catholic church, respectively. The rest of the table resembles the main identification strategy. The independent variable is exposure to conflict. The results show that individuals exposed to conflict do not have different trust in institutions than comparable individuals who were not exposed to conflict. These results could be explained due to the lack of government intervention during the first years of the conflict.

5.2. Own-Indigenous group violence

The main difference between the Shining Path violence and the setting proposed by Rohner et al. (2013a,b) is who committed the violent events. In Rohner et al. (2013a,b) violence was perpetrated by a different ethnic group than your own, which decreased generalized trust and increased interpersonal and ethnic identity. In the Peruvian context, as explained in section 2, violence was perpetrated by younger generations of the same district that joined the Shining path. Therefore, if own-Indigenous group violence drives the results, we expect the nega-

tive association between exposure to violence and negative Indigenous identity is more salient in Indigenous homogeneous districts.

To test this mechanism, we construct a measure of Indigenous diversity at the district level following Alesina et al. (1999). This index measures the probability that two randomly drawn people from the same district belong to different Indigenous groups. A higher value of the index represents a higher probability of encountering another individual of a different Indigenous group.

The idea of using the Indigenous diversity index is that in more homogeneous districts, the likelihood of suffering own-Indigenous-group violence is higher, leading to a lower Indigenous identification. To create Indigenous diversity measures at the district level, we use the ENAHO survey and only individuals born before 1950. We use these people because they were at least 32 years at the time of the first attack of the Shining Path, hence, their ethnic identity was already formed. Further, we use the self-identification measure to construct the index. Last, we create a dummy variable from the Indigenous diversity index, which takes the value of 1 if the Indigenous diversity index is larger than 0, hence more than one Indigenous group the district, and takes the value of 0 if there is only one Indigenous group in the district.

To test the own-Indigenous group violence mechanism, we use the regression specification embedded in equation 1 and interact the main explanatory variable, exposure to violence, with the measure of ethnic diversity, either as a continuous variable or a dummy. Note that in this specification, the baseline group are individuals not exposed to violence in fully homogeneous districts, hence, only one Indigenous group was present in the district.

Table 4 shows the results using the Indigenous diversity index. In columns (1) and (2), we include the Indigenous diversity index dummy, and the dependent variables are whether an Individuals speak an Indigenous language as a mother tongue (column 1) or whether the individual identifies as Indigenous (column 2). In columns (3) and (4), we include the Indigenous diversity index continuous, and the dependent variables are whether an Individuals speak an Indigenous language as a mother tongue (column 3) or whether the individual identifies as Indigenous (column 4). In all specifications we include the full set of fixed effects, district fixed effects, year-interview fixed effects, district-birth fixed effects, and year-birth fixed effects. Further, we include individual controls, sex, age, and age square.

Table 4. Own-Indigenous group violence

	Mother tongue	Identify indigenous	Mother tongue	Identify indigenous
	(1)	(2)	(3)	(4)
Exposure Violence x T	−0.069***	−0.045**	−0.058***	−0.029***
	(0.022)	(0.017)	(0.019)	(0.011)
Exposure Violence x T x Indigenous diversity *dummy*	0.061***	0.034		
	(0.023)	(0.024)		
Exposure Violence x T x Indigenous diversity index			0.213***	0.042
			(0.057)	(0.053)
District-interview FE	✓	✓	✓	✓
District-birth FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
SE Cluster	district-year	district-year	district-year	district-year
N	29,911	25,566	29,911	25,566
Adjusted R^2	0.622	0.472	0.622	0.472

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. In columns (1) and (2), the explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19 interacted with the Indigenous diversity dummy. In columns (3) and (4), the explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19 interacted with the Indigenous diversity index. In columns (1) and (3), the dependent variable is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (2) and (4), the dependent variable is whether an individual identifies Indigenous or not. In all specifications, we restrict the sample to the districts in the border sample. We include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

The results show that in Indigenous homogeneous districts, individuals exposed to violence during the ethnic identity formation years are between 6.9 pp. and 5.8 pp. less likely to speak an indigenous language as their mother tongue and between 4.5 pp. and 2.9 pp. less likely to identify as Indigenous than individuals not exposed to violence during their ethnic identity formation years, respect to the difference in ethnic identification between individuals that were and were not exposed to violence after ethnic identity was formed.

Further, the results show that individuals exposed to violence during their youth in districts with more than one Indigenous group are 6.1 pp. more likely to speak an Indigenous language and 3.4 pp. more likely to identify as indigenous than individuals not exposed to violence in their identity formation years respect to the difference in ethnic identification between individuals that were and were not exposed to violence after ethnic identity was formed. This translates into a 1sd (0.17 units) increase in the Indigenous diversity index is associated with an increase of 5.8 pp. in the likelihood that an individual speaks an Indigenous language as mother tongue and 1.1 pp. increase in the probability that an individual identifies as Indigenous.

These results show an increasing number of Indigenous groups in a municipality, hence, a higher probability that the perpetrator of violence was an Indigenous person from a different Indigenous group increases Indigenous identity. On the other hand, in Indigenous homogeneous communities, the likelihood that the perpetrator is from

the same Indigenous group decreases the probability that an individual identifies as Indigenous—suggesting that own-Indigenous group violence could be the mechanism driving the main results.

6. Conclusion

In this paper, we examine how exposure to conflict violence shapes ethnic identity, drawing on the Peruvian conflict led by the Shining Path as a natural experiment. Beginning in the early 1980s, this insurgency sought to impose a class-based identity, pressuring rural communities to abandon their traditional ethnic identities. Although not explicitly targeted, the Indigenous populations disproportionately endured the conflict's atrocities.

Our analysis leverages individual-level survey data on ethnic identity in combination with detailed event-level data. These datasets enable us to construct a measure of exposure to violence during critical formative years of ethnic identity (0–19) versus later stages (20+). With this data at our disposal, we implement a difference-in-differences strategy that compares changes in ethnic self-identification—both self-ascribed and mother tongue—across cohorts exposed to violence before and after identity formation. By developing a border sample design in which we focus on districts that share similar socio-economic characteristics but differ in their intensity of Shining Path violence, we identify the causal impact of violence conflict on ethnic identity choices.

Our findings indicate that individuals exposed to violence at younger ages are significantly less likely to identify as Indigenous or to speak an Indigenous language as their mother tongue when compared to those who were not exposed. We confirm the parallel trends assumption by showing that cohorts whose identity was largely established, in both violent and non-violent areas, exhibited no meaningful differences in ethnic self-identification. To ensure robustness, we explore whether non-Indigenous violence, selective migration, or alternate dimensions of identity (e.g., class identity) confound these findings. None of these checks challenge our conclusions.

We explore a plausible factor that could explain our main results. We argue that since the main perpetrators were individuals from the same village as the victims, own-ethnic group violence appear to be the main factor driving the results. First, we present evidence that locally perpetrated violence eroded trust. We then demonstrate that in Indigenous homogeneous districts, individuals exposed to violence during their years of identity formation were less likely to retain their Indigenous identity. This contrasts with conflicts involving external aggressors, which can sometimes foster ethnic cohesion.

Taken together, these results highlight the fragility of ethnic identity

in the face of coercive forces and underscore the importance of conflict context: when violence originates within the same ethnic group, individuals may respond by distancing themselves from that identity. These findings contribute to a growing literature on the long-term consequences of violence on social and economic outcomes, offering new evidence on how conflicts can reshape not just physical and economic realities but also fundamental aspects of identity. By examining a historical episode of insurgency in Peru, our study demonstrates the lasting influence that conflict violence can exert on ethnic identity and sheds light on the broader implications for societies in which ethnic identities are both politically salient and vulnerable to external pressures.

References

Ainsworth, Mary S (1989). Attachments beyond infancy. *American psychologist* 44(4), p. 709.

Akerlof, George A and Rachel E Kranton (2000). Economics and identity. *The quarterly journal of economics* 115(3), pp. 715–753.

— (2002). Identity and schooling: Some lessons for the economics of education. *Journal of economic literature* 40(4), pp. 1167–1201.

— (2010). Identity economics. Princeton University Press.

Albertus, Michael (2020). Land reform and civil conflict: Theory and evidence from Peru. *American Journal of Political Science* 64(2), pp. 256–274.

Albertus, Michael, Thomas Brambor, and Ricardo Ceneviva (2018). Land inequality and rural unrest: theory and evidence from Brazil. *Journal of Conflict Resolution* 62(3), pp. 557–596.

Albertus, Michael and Oliver Kaplan (2013). Land reform as a counterinsurgency policy: Evidence from Colombia. *Journal of Conflict Resolution* 57(2), pp. 198–231.

Alesina, Alberto, Reza Baqir, and William Easterly (1999). Public goods and ethnic divisions. *The Quarterly journal of economics* 114(4), pp. 1243–1284.

Arbatlı, Cemal Eren, Quamrul H Ashraf, Oded Galor, and Marc Klemp (2020). Diversity and conflict. *Econometrica* 88(2), pp. 727–797.

Aspachs-Bracons, Oriol, Irma Clots-Figueras, Joan Costa-Font, and Paolo Masella (2008). Compulsory language educational policies and identity formation. *Journal of the European Economic Association* 6(2-3), pp. 434–444.

Basu, Sukanya and Sarah Pearlman (2017). Violence and migration: evidence from Mexico's drug war. *IZA Journal of Development and Migration* 7(1), pp. 1–29.

Battu, Harminder and Yves Zenou (2010). *Oppositional identities and employment for ethnic minorities: Evidence from England*.

Bénabou, Roland and Jean Tirole (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics* 126(2), pp. 805–855.

Bowlby, John (1982). Attachment and loss: retrospect and prospect. *American journal of Orthopsychiatry* 52(4), p. 664.

Brand, Elaine S, Rene A Ruiz, and Amado M Padilla (1974). Ethnic identification and preference: A review. *Psychological Bulletin* 81(11), p. 860.

Cama, Shireen F and Priya Sehgal (2021). Racial and ethnic considerations across child and adolescent development. *Academic Psychiatry* 45(1), pp. 106–109.

Caselli, Francesco and Wilbur John Coleman (2013). On the theory of ethnic conflict. *Journal of the European Economic Association* 11(suppl_1), pp. 161–192.

Casey, Teresa and Christian Dustmann (2010). *Immigrants' identity, economic outcomes and the transmission of identity across generations*.

Chen, Yan and Sherry Xin Li (2009). Group identity and social preferences. *American Economic Review* 99(1), pp. 431–57.

Clots-Figueras, Irma and Paolo Masella (2013). Education, language and identity. *The Economic Journal* 123(570), F332–F357.

Collier, Paul and Anke Hoeffler (2004). Greed and grievance in civil war. *Oxford economic papers* 56(4), pp. 563–595.

Collier, Paul, Anke Hoeffler, and Dominic Rohner (2009). Beyond greed and grievance: feasibility and civil war. *oxford Economic papers* 61(1), pp. 1–27.

Constant, Amelie F and Klaus F Zimmermann (2008). Measuring ethnic identity and its impact on economic behavior. *Journal of the European Economic Association* 6(2-3), pp. 424–433.

CVR, Comisión de la Verdad y Reconciliación (2003). *Informe final de la Comisión de la Verdad y Reconciliación*. URL: https://www.cverdad.org.pe/ifinal/index.php (visited on 05/12/2022).

Davey, AG and PN Mullin (1980). Ethnic identification and preference of British primary school children. *Child Psychology & Psychiatry & Allied Disciplines*.

De Vries, Leonie Ansems and Elspeth Guild (2019). Seeking refuge in Europe: spaces of transit and the violence of migration management. *Journal of ethnic and migration studies* 45(12), pp. 2156–2166.

Erikson, Erik H (1968). Identity: Youth and crisis. 7. WW Norton & company.

Esteban, Joan, Laura Mayoral, and Debraj Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review* 102(4), pp. 1310–42.

Esteban, Joan and Debraj Ray (2011). Linking conflict to inequality and polarization. *American Economic Review* 101(4), pp. 1345–74.

Finkel, Evgeny, Scott Gehlbach, and Tricia D Olsen (2015). Does reform prevent rebellion? Evidence from Russia's emancipation of the serfs. *Comparative Political Studies* 48(8), pp. 984–1019.

Galdo, Jose (2013). The long-run labor-market consequences of civil war: Evidence from the Shining Path in Peru. *Economic Development and Cultural Change* 61(4), pp. 789–823.

García-Ponce, Omar (2017). Women's Political Participation After Civil War: Evidence from Peru. *Manuscript. University of California, Davis. At here. pdf, accessed July* 16, p. 2019.

Guardado, Jenny (2018). Land tenure, price shocks, and insurgency: Evidence from Peru and Colombia. *World Development* 111, pp. 256–269.

Gúzman, Abigael (1974). *LA PROBLEMATICA NACIONAL*. URL: https://cedema.org/digital_items/631 (visited on 05/12/2022).

Hargreaves Heap, Shaun P and Daniel John Zizzo (2009). The value of groups. *American Economic Review* 99(1), pp. 295–323.

Holcomb-McCoy, Cheryl (2005). Ethnic identity development in early adolescence: Implications and recommendations for middle school counselors. *Professional School Counseling* 9(2).

Huntington, Samuel P (2006). Political order in changing societies. Yale university press.

Jima-González, Alexandra and Miguel Paradela-López (2020). Indians in Pensamiento Gonzalo: The Influence of 20th-Century Peruvian Intelligentsia on Shining Path's Ideology. *SAGE Open* 10(4), p. 2158244020982990.

— (2021). The Negative Impact of Shining Path on Indigenous Mobilization in Peru: An Approach from Political Opportunity and New Social Movements Theories. *Latin American Perspectives* 48(6), pp. 194–209.

Kalyvas, Stathis N (2006). The logic of violence in civil war. Cambridge University Press.

Kapstein, Ethan B (2017). Seeds of stability: land reform and US foreign policy. Cambridge University Press.

Kent, Robert B (1993). Geographical dimensions of the Shining Path insurgency in Peru. *Geographical Review*, pp. 441–454.

Kranton, Rachel E (2016). Identity economics 2016: Where do social distinctions and norms come from? *American Economic Review* 106(5), pp. 405–09.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert Vishny (1999). The quality of government. *The Journal of Law, Economics, and Organization* 15(1), pp. 222–279.

Laplante, Lisa J and Kimberly Theidon (2007). Truth with consequences: Justice and reparations in post-Truth Commission Peru. *Hum. Rts. Q.* 29, p. 228.

Leon, Gianmarco (2012). Civil conflict and human capital accumulation the long-term effects of political violence in perú. *Journal of Human Resources* 47(4), pp. 991–1022.

Lichter, Andreas, Max Löffler, and Sebastian Siegloch (2021). The long-term costs of government surveillance: Insights from stasi spying in East Germany. *Journal of the European Economic Association* 19(2), pp. 741–789.

Marcia, James E et al. (1980). Identity in adolescence. *Handbook of adolescent psychology* 9(11), pp. 159–187.

Marks, Amy Kerivan, Laura A Szalacha, Meaghan Lamarre, Michelle J Boyd, and Cynthia García Coll (2007). Emerging ethnic identity and interethnic group social preferences in middle childhood: Findings from the Children of Immigrants Development in Context (CIDC) study. *International Journal of Behavioral Development* 31(5), pp. 501–513.

Mason, T David (1998). " Take Two Acres and Call Me in the Morning": Is Land Reform a Prescription for Peasant Unrest? *The Journal of Politics* 60(1), pp. 199–230.

Montalvo, José G and Marta Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *American economic review* 95(3), pp. 796–816.

Morrison, Andrew R (1993). Violence or economics: what drives internal migration in Guatemala? *Economic Development and Cultural Change* 41(4), pp. 817–831.

Parkes, Colin Murray, Joan Stevenson-Hinde, and Peter Marris (2006). Attachment across the life cycle. Routledge.

PCP, Partido Comunista del Perú [Sendero Luminoso] (1976). *El problema campesino y la revolución*. URL: https://cedema.org/digital_items/635 (visited on 05/12/2022).

Phinney, Jean S (1989). Stages of ethnic identity development in minority group adolescents. *The Journal of Early Adolescence* 9(1-2), pp. 34–49.

— (1990). Ethnic identity in adolescents and adults: review of research. *Psychological bulletin* 108(3), p. 499.

Rohner, Dominic (2011). Reputation, group structure and social tensions. *Journal of Development Economics* 96(2), pp. 188–199.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti (2013a). Seeds of distrust: Conflict in Uganda. *Journal of Economic Growth* 18(3), pp. 217–252.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti (2013b). War signals: A theory of trade, trust, and conflict. *Review of Economic Studies* 80(3), pp. 1114–1147.

Schüller, Simone (2015). Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics* 28(4), pp. 965–1004.

Shining Path (1988). *Documentos fundamentales [Fundamental Documents]*. URL: http://www.solrojo.org/pcp_doc/pcp_gd88.htm (visited on 05/12/2022).

Silva, Adriana Carolina and Douglas S Massey (2015). Violence, networks, and international migration from Colombia. *International Migration* 53(5), pp. 162–178.

Tajfel, Henri and John C Turner (2004). The social identity theory of intergroup behavior. In: *Political psychology*. Psychology Press, pp. 276–293.

Wheat, Andrew (1990). Shining Path's" fourth sword" ideology. *Journal of Political & Military Sociology*, pp. 41–55.

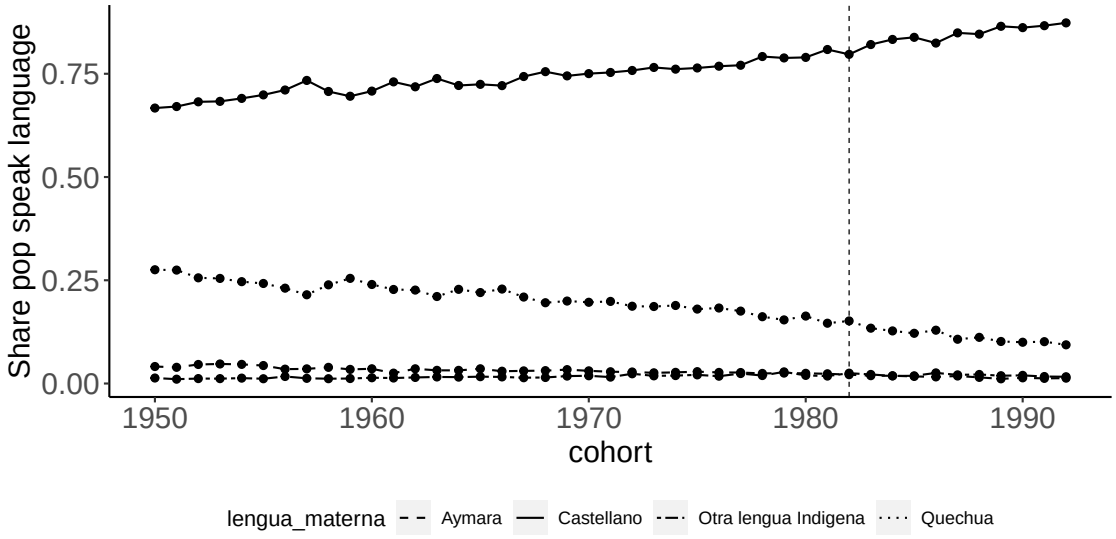
Wood, Elisabeth Jean (2003). Insurgent collective action and civil war in El Salvador. Cambridge University Press.

Xin Li, Sherry (2010). Social identities, ethnic diversity, and tax morale. *Public Finance Review* 38(2), pp. 146–177.

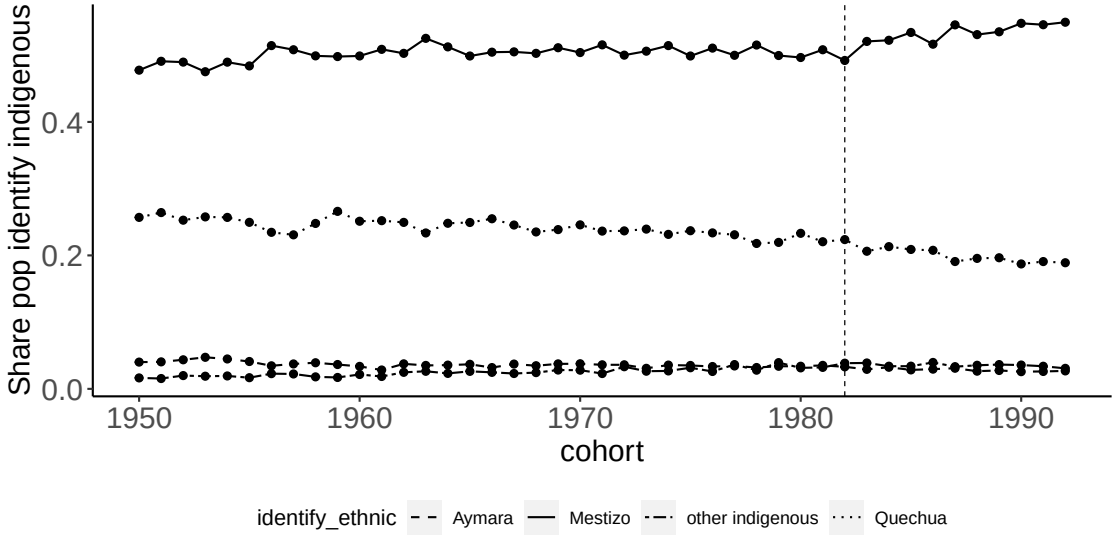
Yip, Tiffany (2014). Ethnic identity in everyday life: The influence of identity development status. *Child Development* 85(1), pp. 205–219.

Appendices

A. Tables and Figures



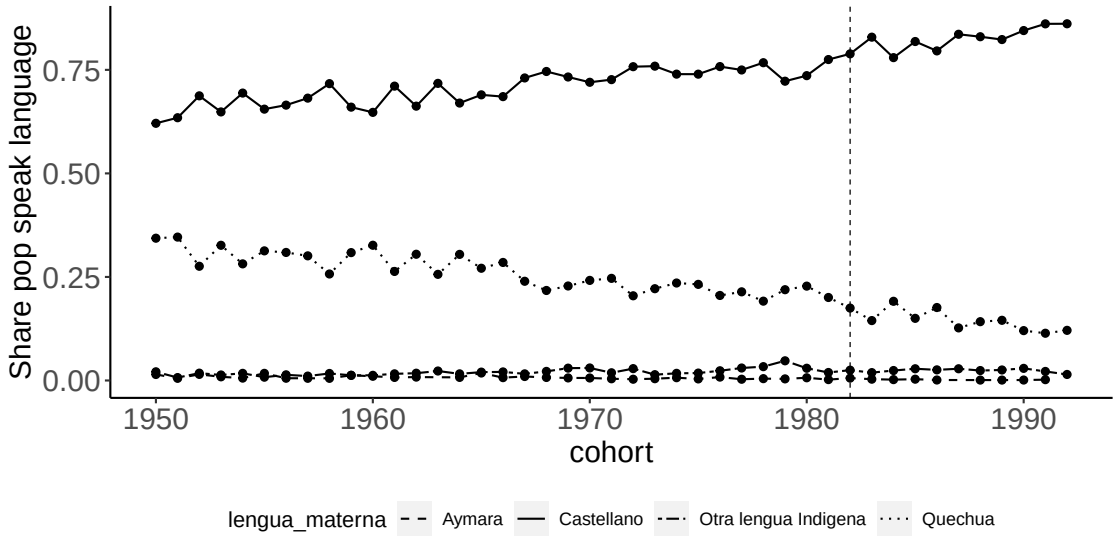
(a) Indigenous language



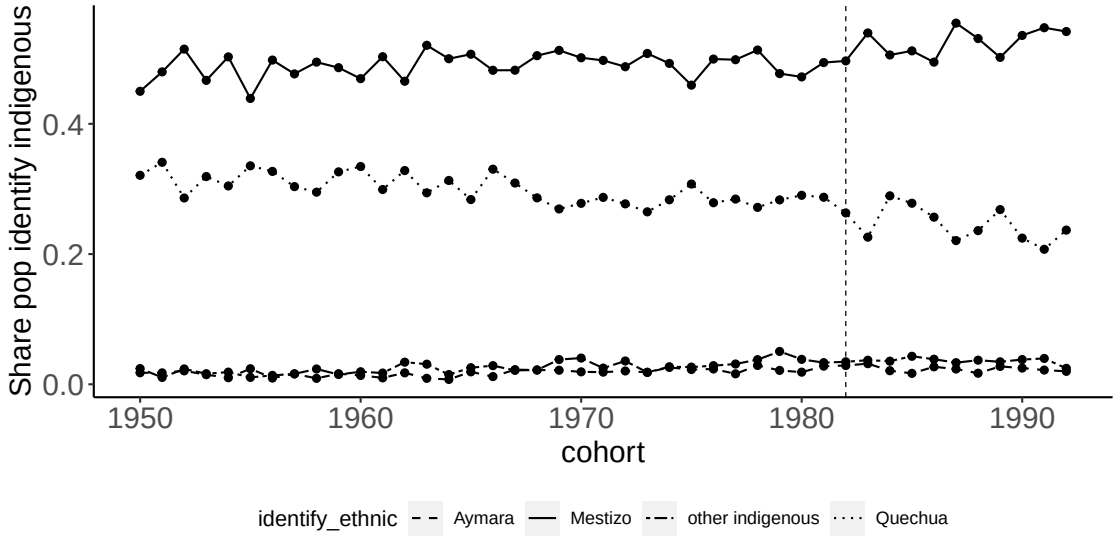
(b) Identify Indigenous

Figure A.1. Ethnic identification: full sample

Notes: This figure plots the share of individuals in each cohort by ethnicity. Subfigure A.1a measures ethnicity as which language individuals speak as their mother tongue. Subfigure A.1b measures ethnicity as which ethnic group the individual self-identify as. The data was collected from the ENAHO surveys Peru from 2013 to 2017.



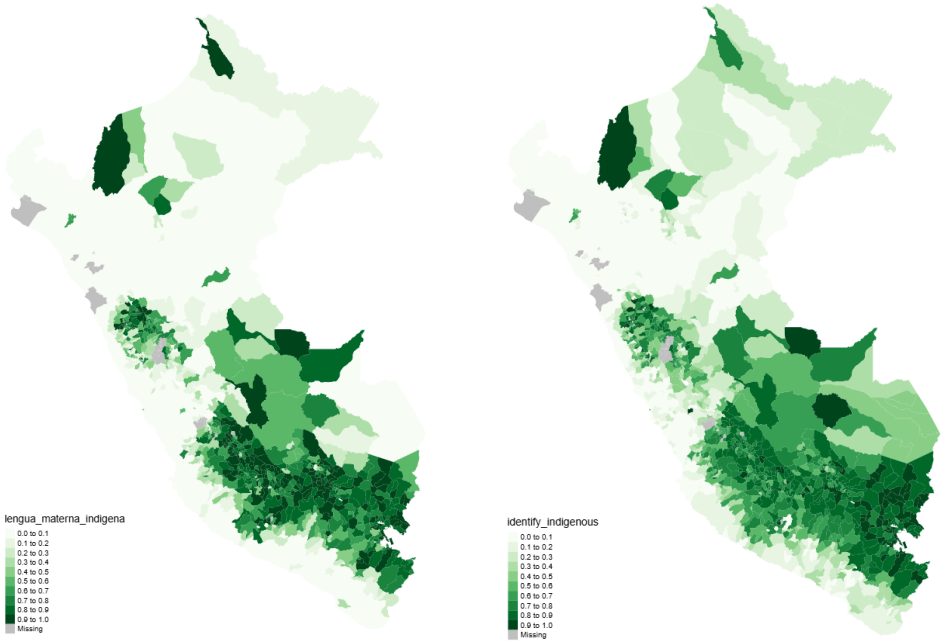
(a) Indigenous language



(b) Identify Indigenous

Figure A.2. Ethnic identification: Border sample

Notes: This figure plots the share of individuals in each cohort by ethnicity. I restrict the sample to the bordering sample. Subfigure A.2a measures ethnicity as which language individuals speak as their mother tongue. Subfigure A.2b measures ethnicity as which ethnic group the individual self-identify as. The data was collected from the ENAHO surveys Peru from 2013 to 2017.

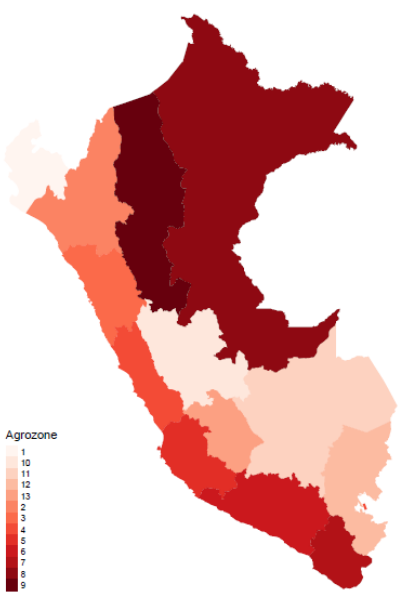


(a) Indigenous language

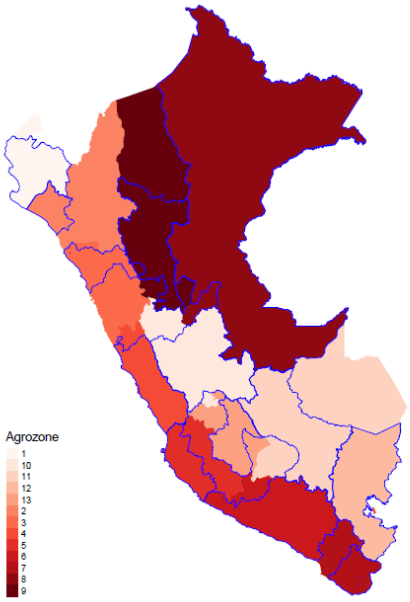
(b) Identify Indigenous

Figure A.3. Indigenous Identity

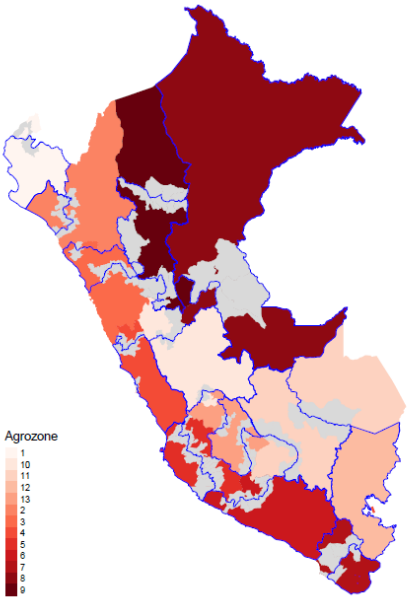
Notes: This figure shows the spatial distribution of the Indigenous people in Peru. Subfigure A.3b plots the share of individuals that speak an Indigenous language as their mother tongue in each Peruvian district. Subfigure A.3a plots the share of individuals that identify indigenous in each Peruvian district. The data was collected from the ENAHO surveys Peru from 2013 to 2017.



(a) Agrozones



(b) Agrozones - core and periphery



(c) Agrozones - core and periphery - sample

Figure A.4. Agrozones

Notes: This figure shows a graphical description of the bordering discontinuity design. Panel A.4a shows Peruvian territory divided by the 13 agrozones, each color represents one agrozone. Panel A.4b adds to the agrozones the borders between "core" and "periphery" districts. The blue line depicts the bordering design. panel A.4c shows in grey all the bordering districts selected for the analysis.

B. Tables

Table A.1. Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Full Sample					
Identify Indigenous	204,247	0.36	0.48	0	1
Indigenous mother tongue	253,332	0.22	0.41	0	1
Exposure violence < 20	254,301	0.17	0.38	0	1
Exposure violence	254,301	0.25	0.44	0	1
Age	254,301	38.95	10.11	20	64
Year born	254,301	1,976	10.00	1,958	1,992
Sex (Female = 1)	254,301	0.52	0.50	0	1
Migration	254,301	0.70	0.46	0	1
Border Sample					
Identify Indigenous	27,842	0.40	0.49	0	1
Indigenous mother tongue	33,326	0.24	0.43	0	1
Exposure violence < 20	33,449	0.27	0.44	0	1
Exposure	33,449	0.36	0.48	0	1
Exposure non-Indigenous violence < 20	33,449	0.25	0.43	0	1
Exposure non-Indigenous violence	33,449	0.35	0.48	0	1
Age	33,449	38.98	10.14	20	59
Year Born	33,449	1,976	10.01	1,958	1,992
Sex (Female = 1)	33,449	0.52	0.50	0	1
Migration	33,449	0.65	0.48	0	1
Indigenous diversity index	30,018	0.10	0.17	0.00	0.49
Indigenous diversity index dummy	30,018	0.42	0.49	0	1
Border Sample variables					
Violent events	229	21.56	83.19	0	661
Deaths	229	40.00	168.46	0	1,981
Vote Marxist	229	0.12	0.12	0.00	0.88
Non-Indigenous population 1961	229	0.59	0.40	0.003	1.00
Illiteracy rate 1961	229	0.53	0.22	0.01	1.00
Illiteracy rate 1972	229	0.38	0.15	0.08	0.68
Slope	229	6.70	3.58	0.07	18.23
Social movement	229	0.17	0.77	0	6
Share cultivated land	229	5.25	5.58	0.00	37.19
Hectares private land	229	23.16	177.98	0.00	2,680.50
Elevation	229	2.55	1.28	0.04	4.60
State employees 1961	229	2.74	1.45	0.00	8.37
Population 1972 (thousands)	229	8.44	22.16	0.22	230.81
log Land reform hectares	229	4.23	10.30	0.00	81.45

Notes: The table shows summary statistic for all variables included in the empirical analysis.

Table A.2. Total: District SE

	Indigenous language as mother tongue		Identify Indigenous	
	all districts	border sample	all districts	border sample
	(1)	(2)	(3)	(4)
Exposure violence x T	−0.014**	−0.034*	−0.005	−0.023**
	(0.007)	(0.018)	(0.005)	(0.011)
District-interview FE	✓	✓	✓	✓
District-born FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DP	0.23	0.25	0.36	0.40
SE Cluster	district	district	district	district
N	208,768	33,326	168,301	27,842
Adjusted R^2	0.660	0.628	0.545	0.480

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1)-(2) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (3)-(4), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (3), we use the full sample of individuals. In columns (2) and (4), we restrict the sample to the districts in the border sample. In all specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.3. Total: No post cohort

	Indigenous language as mother tongue		Identify Indigenous	
	all districts	border sample	all districts	border sample
	(1)	(2)	(3)	(4)
Exposure violence x T	−0.051***	−0.046	−0.009	−0.022*
	(0.010)	(0.030)	(0.006)	(0.013)
	(0.006)	(0.016)	(0.005)	(0.009)
District-interview FE	✓	✓	✓	✓
District-born FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DP	0.23	0.25	0.36	0.40
SE Cluster	district-year	district-year	district-year	district-year
N	204,434	32,417	164,565	27,048
Adjusted R^2	0.666	0.628	0.548	0.477

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1)-(2) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (3)-(4), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (3), we use the full sample of individuals except for the post cohort. In columns (2) and (4), we restrict the sample to the districts in the border sample. In all specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.4. Migration

	Indigenous Language as mother tongue			identify Indigenous			Migration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure violence x T	−0.034**	−0.034**	−0.050	−0.023**	−0.023**	−0.035**		0.006
	(0.016)	(0.016)	(0.037)	(0.009)	(0.009)	(0.014)		(0.004)
Migration		0.007	0.001		0.008	0.003		
		(0.004)	(0.008)		(0.008)	(0.011)		
Migration x Exposure			0.004			0.008		
			(0.017)			(0.020)		
Exposure violence x T x Migration			0.025			0.018		
			(0.039)			(0.018)		
Exposure							0.084***	
							(0.002)	
District-interview FE	✓	✓	✓	✓	✓	✓		✓
District-birth FE	✓	✓	✓	✓	✓	✓		✓
Year-interview FE	✓	✓	✓	✓	✓	✓		✓
Year-birth FE	✓	✓	✓	✓	✓	✓		✓
Controls	✓	✓	✓	✓	✓	✓		✓
SE Cluster	district-year	district-year	district-year	district-year	district-year	district-year		district-year
Sample	Border	Border	Border	Border	Border	Border	Full	Border
N	33,326	33,326	33,326	27,842	27,842	27,842	254,301	33,449
Adjusted R^2	0.628	0.628	0.628	0.480	0.480	0.480	0.006	0.829

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. In Column (7), explanatory variables is whether an individual was exposed to violence or not and we use the full sample in this column. The dependent variable in columns (1)-(3) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (4)-(6), the dependent variable is whether an individual identifies Indigenous or not. In columns (7) and (8), the dependent variable is whether an individual migrated or not. In columns (2) and (5), we include a dummy variable indicating whether an individual migrated or not as a control variable. In columns (3) and (6), we interact the main explanatory variable with the *migration* dummy. In all specifications except for column (7), we restrict the sample to the districts in the border sample. We include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.5. Ethnic groups

	Indigenous language as mother tongue			Identify Indigenous		
	Quechua	Aymara	Other Indigenous	Quechua	Aymara	Other Indigenous
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure violence x T	−0.031**	0.002	−0.007*	−0.011	−0.007	−0.012**
	(0.015)	(0.002)	(0.004)	(0.007)	(0.007)	(0.005)
Intensive margin controls	✓	✓	✓	✓	✓	✓
District-interview FE	✓	✓	✓	✓	✓	✓
District-birth FE	✓	✓	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean DP	0.22	0.006	0.029	0.22	0.006	0.0294
SE Cluster	district-year	district-year	district-year	district-year	district-year	district-year
N	32,395	25,471	26,074	26,130	17,509	17,869
Adjusted R^2	0.660	0.135	0.658	0.559	0.255	0.500

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1)-(3) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (4)-(6), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (4), we only include the Quechua population. In columns (2) and (5), we only include the Aymara population. In columns (3) and (6), we only include Other Indigenous population. In all specifications, we restrict the sample to the districts in the border sample. We include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.6. Total

	Indigenous language as mother tongue		Identify Indigenous	
	all districts	border sample	all districts	border sample
	(1)	(2)	(3)	(4)
Exposure cohort post	−0.130***	−0.034	−0.017	−0.001
	(0.017)	(0.029)	(0.031)	(0.031)
Exposure cohort 0-4	−0.088***	−0.070***	−0.020	−0.040**
	(0.010)	(0.023)	(0.018)	(0.018)
Exposure cohort 5-9	−0.049***	−0.044**	−0.007	−0.025
	(0.008)	(0.019)	(0.020)	(0.020)
Exposure cohort 10-14	−0.045***	−0.029	−0.014	−0.017
	(0.006)	(0.022)	(0.020)	(0.020)
Exposure cohort 15-19	−0.014***	−0.006	0.009	0.008
	(0.005)	(0.016)	(0.013)	(0.013)
Exposure cohort 25-29	0.007	0.036*	0.004	0.023
	(0.011)	(0.021)	(0.024)	(0.024)
Exposure cohort 30-34	0.003	−0.015	0.012	0.020
	(0.017)	(0.060)	(0.035)	(0.035)
District-interview FE	✓	✓	✓	✓
District-birth FE	✓	✓	✓	✓
Year-interview FE	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean DP	0.23	0.25	0.36	0.40
N	208,768	33,326	168,301	27,842
Adjusted R^2	0.661	0.629	0.545	0.480

Notes: The table shows results from estimating equation 2. The unit of observation is the individual. The omitted cohort is 20-24. The dependent variable in columns (1)-(2) is whether an individual speaks an Indigenous language as their mother tongue or not. In columns (3)-(4), the dependent variable is whether an individual identifies Indigenous or not. In columns (1) and (3), we use the full sample of individuals. In columns (2) and (4), we restrict the sample to the districts in the border sample. In all specifications, we include the full set of fixed effects, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.7. Total

	totalevents	deaths	marxistleft	nonindigenous_pc_1961	illiterate_pc_1961	illiterate_pc_1972	slope
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
inside_agrozone_core	30.004***	53.254**	0.021	−0.067	−0.026	−0.021	−0.977**
	(10.868)	(22.093)	(0.016)	(0.053)	(0.029)	(0.019)	(0.471)
N	229	229	229	229	229	229	229
R^2	0.032	0.025	0.007	0.007	0.003	0.005	0.019

	mov_comunidad	cult_land	prop_ha_ths	elev_1k	lemployees1961	population_1972ths	LRpercap_calweighted_log
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
inside_agrozone_core	0.151	0.533	22.560	−0.428**	0.119	3.861	1.355
	(0.101)	(0.740)	(23.592)	(0.168)	(0.193)	(2.932)	(1.365)
N	229	229	229	229	229	229	229
R^2	0.010	0.002	0.004	0.028	0.002	0.008	0.004

Notes: The table shows the standardized mean differences of different variables between core and periphery districts. The base category is core districts. The unit of observation is the districts. The variables are total number of violent events, total number of death, share of votes to Marxist parties , share of non-Indigenous population in 1961, illiteracy rate in 1961, , illiteracy rate in 1972, peasant movements, average slope, share of cultivated land, Hectares of private land, average elevation, number of state employees, population in 1972 (thousands), Hectares of land reform. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.8. Trust interpersonal

	Interpersonal trust	
	(1)	(2)
Exposure violence x T	−0.167*	−0.147
	(0.101)	(0.095)
No-Indigenous		0.109**
		(0.052)
District FE	✓	✓
Survey FE	✓	✓
Year-birth FE	✓	✓
Controls	✓	✓
Mean DV	2.46	2.46
SE Cluster	province-year	province-year
N	5,158	5,039
Adjusted R^2	0.032	0.033

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable is the measure of interpersonal trust. In column (2), we include a dummy variable indicating whether the individual is Indigenous or not. In all specifications, we include, year-birth fixed effects, year-survey fixed effects, district fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.9. Trust: Generalized

	Generalized trust	
	(1)	(2)
Exposure violence x T	0.014	0.013
	(0.018)	(0.017)
No-Indigenous		−0.010
		(0.030)
District FE	✓	✓
Survey FE	✓	✓
Year-birth FE	✓	✓
Controls	✓	✓
Mean DV	0.16	0.16
SE Cluster	province-year	province-year
N	5,905	5,731
Adjusted R^2	0.018	0.011

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable is the measure of generalized trust. In column (2), we include a dummy variable indicating whether the individual is Indigenous or not. In all specifications, we include, year-birth fixed effects, year-survey fixed effects, district fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.10. Trust: Institutions

	district	national	judiciary	congress	know democracy	catholic church
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure violence x T	0.009	−0.003	0.035	−0.013	−0.005	0.012
	(0.020)	(0.015)	(0.022)	(0.020)	(0.007)	(0.025)
District-interview FE	✓	✓	✓	✓	✓	✓
District-birthFE	✓	✓	✓	✓	✓	✓
Agrozone FE	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean DP	1.88	1.74	1.78	1.60	0.53	2.49
SE Cluster	province-year	province-year	province-year	province-year	province-year	province-year
N	28,676	27,471	27,860	27,603	29,804	29,168
Adjusted R^2	0.056	0.048	0.048	0.058	0.182	0.084

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. We restrict the sample to the bordering sample. The dependent variable in columns (1),(2), (3), (4), (5), (6) are measures of trust in the district government, the national government, the judiciary power, the congress, and the catholic church. In all specifications, we include, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

Table A.11. Trust: Institutions (Full sample)

	district	national	judiciary	congress	know democracy	catholic church
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure violence	0.011*	−0.012	0.011*	−0.002	−0.007*	0.004
	(0.006)	(0.011)	(0.006)	(0.014)	(0.004)	(0.012)
District-interview FE	✓	✓	✓	✓	✓	✓
District-birth FE	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓
Year-birth FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean DP	1.87	1.74	1.77	1.60	0.53	2.52
SE Cluster	province-year	province-year	province-year	province-year	province-year	province-year
N	181,013	173,637	176,110	174,345	187,703	183,955
Adjusted R^2	0.056	0.062	0.052	0.059	0.157	0.084

Notes: The table shows results from estimating equation 1. The unit of observation is the individual. The explanatory variable is the interaction term between exposure to violence and exposure to violence up to age of 19. The dependent variable in columns (1),(2), (3), (4), (5), (6) are measures of trust in the district government, the national government, the judiciary power, the congress, and the catholic church. In all specifications, we include, year-birth fixed effects, year-interview fixed effects, district of interview fixed effects, and district of birth fixed effects. Baseline controls include age, age square, and sex. Standard errors clustered at the district-year level in parenthesis. Statistical significance at the 99%, 95%, and 90% confidence level denoted: ***, **, and *.

C. Description of Variables

C.1. Main variables

Identify Indigenous.

Dummy variable which takes the value of 1 if the individual answered to the question: *"According to your ancestor and your customs, do you consider yourself: Quechua, Aymara, Other Indigenous community, Black, White, Mestizo, or unknown?"*, Quechua, Aymara, Other Indigenous community, and takes the value of 0 if the individual answered Mestizo. We restrict the sample to individuals born between 1958 and 1992. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

Indigenous mother tongue.

Dummy variable which takes the value of 1 if the individual answered to the question: "*What is the language or mother tongue you learned in your childhood*", Quechua, Aymara, Other Indigenous language, and takes the value of 0 if the individual answered Castellano (Spanish). We restrict the sample to individuals born between 1958 and 1992. Data obtained from Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

Exposure to violence.

Dummy variable which takes the value 1 if individual i was born in a district that was exposed to a violent attack of the Shining Path between 1980 and 1992 were the victim was an Indigenous person. We restrict the sample to individuals born between 1958 and 1992. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO), and Truth and Reconciliation Commission (TRC).

Exposure to non-Indigenous violence.

Dummy variable which takes the value 1 if individual i was born in a district that was exposed to a violent attack of the Shining Path between 1980 and 1992 were the victim was not an Indigenous person. We restrict the sample to individuals born between 1958 and 1992. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO), and Truth and Reconciliation Commission (TRC).

Exposure to violence < 20**.**

Dummy variable which takes the value 1 if individual i was born in a district that was exposed to a violent attack of the Shining Path between 1980 and 1992 between age 0 and 19 were the victim was an Indigenous person, and takes value 0 otherwise. We restrict the sample to individuals born between 1958 and 1992. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO). and Truth and Reconciliation Commission (TRC).

Exposure to non-Indigenous violence < 20**.**

Dummy variable which takes the value 1 if individual i was born in a district that was exposed to a violent attack of the Shining Path between 1980 and 1992 between age 0 and 19 were the victim was not an

Indigenous person, and takes value 0 otherwise. We restrict the sample to individuals born between 1958 and 1992. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO). and Truth and Reconciliation Commission (TRC).

Age.

Age of individual i. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

Year born.

Year born of individual i. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

Sex.

Dummy variable which takes value 1 for female, and 0 for male. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

Indigenous diversity index.

Measures the probability that two randomly drawn people from the same district belong to different Indigenous groups. A higher value of the index represents a higher probability of encountering another individual of a different Indigenous group. We restrict the sample to individuals born before 1950. To create the index, we follow Alesina et al. (1999), and use the following equation 3, where $Indigenous_i d$ represent the share of population self-identify from indigenous group $i \in$ (Quechua, Aymara, Other Indigenous) in district d. Data obtained from Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza (ENAHO).

$$Indigenous\ diversity\ index = 1 - \sum_i (Indigenous_i d)^2 \tag{3}$$

Indigenous diversity index dummy.

Dummy variable which takes the value of 1 if the Indigenous diversity index is larger than 0, and takes the value of 0 otherwise.

C.2. Sampling design variables

Total events.

The number of attacks during the Peruvian civil conflict, which include murders, disappearances, illegal detentions, kidnappings, forced recruitment, torture, rape, and woundings, at the district level from 1980 to 2000. Data obtained from Albertus (2020).

Deaths.

The deaths during the Peruvian civil conflict at the district level from 1980 to 2000. Data obtained from Albertus (2020).

Vote Marxist.

Marxist vote share in the 1980 election. Data obtained from Albertus (2020).

Non-Indigenous pop in 1961.

The share of non-Indigenous population in 1961 at the district level. Data obtained from Albertus (2020).

Non-Indigenous pop in 1961.

The share of non-Indigenous population in 1961 at the district level. Data obtained from Albertus (2020).

Non-Indigenous pop in 1961.

Measure as the share of individuals that speak Spanish as their mother tongue in 1961 at the district level. Data obtained from Albertus (2020).

Land reform Hc.

Land reform per capita at the district level. Data obtained from Albertus (2020).

Population in 1972.

The number of inhabitants in thousands in 1972 at the district level. Data obtained from Albertus (2020).

State Employees.

The number of state personnel in 1961 at the district level. Data obtained from Albertus (2020).

Hc private land.

The number hectares of private land in thousands at the district level. Data obtained from Albertus (2020).

Elevation.

The average elevation in meters at the district level. Data obtained from Albertus (2020).

Share cultivable land.

The share of cultivable land at the district level. Data obtained from Albertus (2020).

Social movement.

The number of previous autonomous communal uprisings at the district level. Data obtained from Albertus (2020).

Slope.

The average slope in degree at the district level. Data obtained from Albertus (2020).

Illiteracy rate 1972.

Share of population without reading and writing ability in 1972 at the district level. Data obtained from Albertus (2020).

Illiteracy rate 1961.

Share of population without reading and writing ability in 1972 at the district level. Data obtained from Albertus (2020).

D. Discussion Albertus (2020)

The sample border design in this paper is substantially different than Albertus (2020) study. First, we do a bordering discontinuity design à la Lichter et al. (2021) and select the sample of municipalities at

the land reform zone borders, whereas Albertus (2020) use the exogenous border and perform a regression discontinuity design. Second, we include all the "agrarian reform zones" in Peru, and Albertus (2020) includes "agrarian reform zones" 1 to 7 and 11. Third, we perform a difference-in-differences analysis, which allows us to compare cohorts that were exposed to violence against cohorts that were not exposed to violence in districts with and without conflict events, and Albertus (2020) focuses on the effect of land reform on conflict.

The literature has studied the relationship between land reform and conflict. The evidence regarding the association between land reform and conflict is mixed. Some studies find a positive relationship[11] and others a negative relationship.[12] Remarkably, authors had found opposites results on the association between land reform and violence for same land reform and conflict. For example, in the Colombian land reform, Guardado (2018) finds a negative relationship through the opportunity cost mechanism. The opportunity cost is the most commonly proposed mechanism for explaining the negative relationship between conflict and land reform. Landless peasants have no opportunity cost for entering a guerrilla movement, increasing the likelihood of conflict.[13] On the other hand, Albertus and Kaplan (2013) finds a positive relationship between land reform and conflict through the "low-intensity" land reform mechanism. A "low-intensity" land reform could incentivize recipients of land reform to fight for more land redistribution when land reforms do not provide significant improvements for those who received land, leading to an increase in the probability of conflict. Last, another mechanism explaining the positive relationship between land reform and conflict is the lack of government investment in infrastructures such as irrigation or market access that landowners provided before the reform, leading to a decrease in production and an increase in the likelihood of conflict (Kapstein, 2017).

The main difference between the bordering sample of this paper is that we include all agrozones and Albertus (2020) omitted Zones 8, and 9 because no major land reform applies; Zone 10 because they split into two zones 10, and 13 in 1974; and Zone 12 because it was not split by core and peripheral areas. The inclusion of Zones 8, and 9 could explain the positive relationship through the "low intensity" mechanism. If, in these Zones, land reform was low, it could incentivize peasants to join the Shining Path Guerrilla due to the Shining Path rhetoric of social justice. Furthermore, the inclusion of Zones 10 and 13 could explain the positive relationship between land reform and conflict through the

[11] Positve relationship: Albertus et al. (2018), Albertus and Kaplan (2013), Finkel et al. (2015), and Mason (1998)

[12] Negative relationship: Albertus (2020), Guardado (2018), Huntington (2006), Kalyvas (2006), and Wood (2003)

[13] This mechanism explains Albertus (2020) findings

lack of state capacity mechanism. Zones 10 and 13 are located in the Departments of Junin and Pasco, which are bordering departments to the department of Ayacucho, the crane of the Shining Path guerrilla. As discussed in the background section, the Shining Path propagated fast through rural areas because of the low state presence. Therefore, we hypothesize that the lack of state presence in those Departments could have caused a lack of government infrastructure leading to an increasing conflict in areas with higher land reform.