

Elizalde, Aldo; Hidalgo, Eduardo; Kampanelis, Sotiris

Working Paper

Violence and ethnic identity

QUCHE Working Paper Series, No. 26-04

Provided in Cooperation with:

Queen's University Centre for Economic History (QUCHE), Queen's University Belfast

Suggested Citation: Elizalde, Aldo; Hidalgo, Eduardo; Kampanelis, Sotiris (2026) : Violence and ethnic identity, QUCHE Working Paper Series, No. 26-04, Queen's University Centre for Economic History (QUCHE), Belfast

This Version is available at:

<https://hdl.handle.net/10419/341401>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



QUCHEH WORKING PAPER SERIES
<http://www.quceh.org.uk/working-papers>

VIOLENCE AND ETHNIC IDENTITY

Aldo Elizalde (Queen's University Belfast)
Eduardo Hidalgo (University of Cologne)
Sotiris Kampanelis (Cardiff University)

Working Paper 26-04

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's Business School
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE

May 2026

Violence and Ethnic Identity^{*}

Aldo Elizalde[†]
Queen's University Belfast

Eduardo Hidalgo[‡]
University of Cologne

Sotiris Kampanelis[§]
Cardiff University

May 28, 2026

Abstract

Can violence erase an ethnic identity? We study this question using the Shining Path insurgency in Peru (1980–1992), which sought to replace competing affiliations with a single class identity and killed roughly 70,000 people, three-quarters of them Indigenous. Combining individual-level data on self-identification and mother tongue with event-level data on Shining Path violence, we implement a difference-in-differences design that exploits variation in exposure across cohorts at different stages of identity formation, within a matched border sample of pre-conflict comparable districts. Individuals exposed during their formative years (ages 0–19) are substantially less likely to identify as Indigenous or to speak an Indigenous language than cohorts whose identity was already formed when violence arrived. The effect is concentrated mostly in early childhood and particularly pronounced for mother tongue; the margin most directly shaped by parental transmission. Violence against non-Indigenous victims has no comparable effect. The mechanism is *intra-group* violence: because perpetrators were overwhelmingly co-ethnics from the same villages, visible Indigenous identity offered no protection and imposed costs. Consistent with this, the effect is strongest in Indigenous-homogeneous districts.

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, and Carola Stapper for helpful comments. We would also like to thank the audiences at conferences and workshops in IEA Belfast 2025 and LACEA-EHN Workshop on Historical Development 2025.

[†] Corresponding author. Department of Economics, Queen's University Belfast. a.elizalde@qub.ac.uk

[‡] University of Cologne. eduardoivanvillacis@gmail.com

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

1. Introduction

Can a decade of civil violence leave a measurable mark on the ethnic identity of those exposed to it? This question matters for economics because ethnic identity shapes behavior, cooperation, and economic outcomes (Akerlof and Kranton, 2000, 2010; Chen and Li, 2009), and because ethnic diversity and polarization have been linked to conflict, under-provision of public goods, and weak institutions (Alesina et al., 1999; Esteban et al., 2012; Montalvo and Reynal-Querol, 2005). A smaller but growing literature has begun to trace the economic determinants of ethnic identity itself, documenting effects of education policy (Aspachs-Bracons et al., 2008; Clots-Figueras and Masella, 2013) and social environment (Battu and Zenou, 2010). Yet ethnicity is still typically treated, in empirical work, as a fixed trait. Whether and how violent conflict reshapes ethnic identity, plausibly the most intense shock a society can impose on its members, remains comparatively underexplored.

The evidence that does exist points in one direction. In the leading account of Rohner et al. (2013a,b), violence erodes generalized trust, pushes individuals to retreat into their ethnic group, and hardens ethnic attachment. Societies can become locked into self-sustaining cycles of ethnic conflict. This view rests on a specific implicit assumption about the structure of violence: that it is perpetrated by one group against another. When violence is inter-group, one's co-ethnics serve as a protective in-group, and visible ethnic identity coordinates collective defense.

Less is known about what happens when violence is not inter-ethnic but intra-group, that is, when perpetrators and victims belong to the same community. The distinction matters because the mechanism thought to drive the hardening of identity under inter-group violence has no analogue under intra-group violence. If anything, visible ethnic identity under intra-group violence marks one as embedded in the community from which violence comes. The implication is that intra-group violence should produce the opposite outcome from the standard result: ethnic identity should contract rather than expand.

This paper shows that this is what happens. We study the Shining Path insurgency in Peru (1980 to 1992), a Maoist guerrilla movement that sought to replace every competing affiliation, ethnic, religious, regional, with a single class identity, and enforced this vision through massacre, torture, forced recruitment, and mass displacement. The conflict killed roughly 70,000 people, three-quarters of them Indigenous, though Indigenous Peruvians were not targeted on ethnic grounds. Using this conflict as a natural experiment, we find that individuals exposed to violence during the formative years of ethnic identity formation are significantly less likely to identify as Indigenous or to speak an Indigenous language as adults than comparable individuals whose identity was already formed when violence arrived.

The Shining Path offers three features that make it unusually well-suited for the task. First, the insurgency was ideological rather than ethnic: it targeted peasants who resisted, not Indigenous

people as such. Indigenous Peruvians were the majority of victims because they were the majority of the rural population in the departments where the Shining Path operated, not because their ethnicity was the basis for their selection. This breaks the identification problem that confounds most ethnic conflicts, in which violence and ethnic targeting are one and the same. Second, the perpetrators were overwhelmingly co-ethnics. Shining Path cells recruited locally, often by force, and executions were typically carried out by young cadres drawn from the same districts as their victims. A Lucanamarca survivor recalled the killers as the sons of neighbors, named by their mothers' surnames: "they were from the village, we recognized them."¹ The Truth and Reconciliation Commission's event-level data confirm this pattern directly: 63 percent of Shining Path violence against Indigenous victims occurred in the victim's own birth district, Peru's smallest administrative unit. Third, the conflict generated plausibly exogenous variation. It lasted roughly a decade (1982–1992), reached about a third of Peru's districts, and its spatial pattern, while not random, can be credibly conditioned on using a matched sample of bordering districts similar on pre-conflict observables.

Our empirical strategy exploits this variation together with the fact that ethnic identity is known to form during a specific developmental window. The psychological literature places the formation of ethnic self-concept between roughly ages 4 and 19, with identity largely crystallized thereafter (Erikson, 1968; Phinney, 1989). We combine individual-level data on self-identification and mother tongue from the ENAHO household surveys (2013–2017) with event-level data on Shining Path violence from the Truth and Reconciliation Commission (TRC), matched by district and year to each individual's birthplace and year of birth. We then estimate a difference-in-differences specification that compares Indigenous identity between cohorts exposed to violence during their formative years (ages 0–19) and cohorts exposed only after identity formation (ages 20 and above), within districts drawn from a matched border sample where pre-conflict socioeconomic characteristics are balanced.

The central finding is large, precise, and concentrated in the predicted age window. Individuals exposed to violence during identity formation are 3.4 percentage points less likely to speak an Indigenous language as their mother tongue, and 2.3 percentage points less likely to self-identify as Indigenous, than comparable individuals not so exposed, declines of 13.6% and 5.8% relative to the sample means. Event-study estimates show that the effect is driven entirely by exposure before age ten, and most sharply for mother tongue, the margin most directly shaped by parental transmission: children exposed between ages 0 and 4 are 7.8 percentage points less likely to speak an Indigenous mother tongue, while cohorts exposed between ages 10 and 19 show no detectable effect. Placebo cohorts aged 25–34 at exposure, and cohorts born after violence had ended in their

¹ CVR, *Informe Final*, Tomo VII (La Masacre de Lucanamarca), p. 41, citing Testimonio 200357 (Huancasancos, Santiago de Lucanamarca, February 2002). Original Spanish: "ellos eran del pueblo, los hemos reconocido."

district, show null effects, supporting the identifying assumption of parallel trends.

We show that these results survive a battery of challenges. Violence in which the victim was non-Indigenous has no effect on Indigenous identity, ruling out a generic “exposure to any violence” interpretation. Migration does not confound the estimates. This is because the effect is plausibly driven by those who remained in their birth districts, and exposure to violence does not itself predict migration in our sample. The effect is specific to ethnic identity: class identification is unaffected, while Peruvian national identity declines modestly. Heterogeneity is concentrated among the Quechua majority; estimates for the much smaller Aymara population are imprecise.

We then turn to mechanism. If the result were driven by the same channel that Rohner et al. (2013a) identify in Uganda, namely generalized distrust pushing individuals to retreat into their ethnic group, we would expect violence to strengthen ethnic attachment and reinforce trust within the local community. We find the opposite. Exposure lowers interpersonal trust, leaves generalized and institutional trust largely unchanged, and reduces rather than strengthens Indigenous identity. This pattern is difficult to reconcile with the standard inter-group conflict mechanism, but is consistent with a setting in which the primary source of threat came from within one’s own community rather than outside it.

This leads to our proposed mechanism: the Shining Path’s violence was *intra-group*, and intra-group violence changes the payoff structure of ethnic identity in ways that inter-group violence does not. When an outgroup attacks, co-ethnics are allies, and visible ethnic identity can coordinate protection, trust, and collective defense. When co-ethnics themselves are the perpetrators, visible ethnic identity provides no comparable protection and may instead increase the salience of one’s embeddedness in the very social networks from which violence emerges. Parents, we hypothesize, responded along the margin they could control most directly: the intergenerational transmission of language and custom.

The evidence is consistent with this interpretation. First, as noted above, Shining Path violence against Indigenous victims was overwhelmingly local, placing victims on the same social terrain as those who carried out the violence. Second, the trust response reverses the pattern predicted by the standard inter-group conflict mechanism: interpersonal trust falls, while generalized and institutional trust remain largely unchanged. This suggests that violence eroded trust in the local community rather than producing a generalized collapse of social trust or a retreat into ethnic attachment. Third, and most directly, the identity effect is concentrated where intra-group violence was most likely. We construct a district-level index of Indigenous diversity following Alesina et al. (1999), first using ENAHO respondents born before 1950 and then the 1981 Peruvian census, fixed before the conflict began. In Indigenous-homogeneous districts, exposure during formative years reduces the probability of speaking an Indigenous mother tongue by 5.8 percentage points; this effect weakens substantially as Indigenous diversity rises and survives when diversity is measured

using the pre-conflict census. Taken together, the evidence suggests that ethnic identity proved more resilient where violence could plausibly be interpreted through an outgroup lens, but substantially more fragile where threat emerged from within.

Our paper contributes to three literatures. The first is the economics of identity (Akerlof and Kranton, 2000, 2010; Chen and Li, 2009; Kranton, 2016). This literature has established that identity shapes behavior, but the economic determinants of ethnic identity itself remain comparatively underexplored, with notable exceptions studying education policy (Aspachs-Bracons et al., 2008; Clots-Figueras and Masella, 2013) and social environment (Battu and Zenou, 2010). We add violent conflict—and specifically, its intra-group variant—to this list of forces that shape who people become.

The second is the large literature on ethnicity and conflict (Alesina et al., 1999; Esteban et al., 2012; Esteban and Ray, 2011; Montalvo and Reynal-Querol, 2005), which has mapped how ethnic composition predicts conflict. Our contribution is to reverse the causal arrow: we ask how conflict reshapes ethnicity. Here our closest antecedents are Rohner et al. (2013b) and Rohner et al. (2013a), who show that inter-ethnic conflict in Uganda strengthened ethnic attachment and eroded generalized trust, trapping societies in cycles of violence. Our finding is the mirror image. When violence is intra-group rather than inter-group, ethnic attachment weakens rather than strengthens, which may help explain why the Peruvian conflict, unlike many of the cases in the Rohner et al. framework, did not reproduce itself across generations.

The third is a growing literature on the specific consequences of the Shining Path. Existing work has established effects on land reform and conflict incidence (Albertus, 2020), female political participation (García-Ponce, 2017), labor market outcomes (Galdo, 2013), and human capital (Leon, 2012). We show that the Shining Path’s legacy reaches further still: it changed who Peruvians say they are.

More broadly, our findings speak to a basic question about the resilience of ethnic identity under pressure. A large literature, implicit in much of the economics of identity, treats ethnicity as deep and slow-moving, a legacy of centuries rather than something a decade of violence can dislodge. Our evidence suggests a more contingent view. Ethnic identity is durable when the group can cohere against a common external threat. It is fragile when the threat comes from within.

2. Background

2.1. The Shining Path conflict

The Shining Path (*Sendero Luminoso*) began as a faction of the Peruvian Communist Party (PCP). In 1970, Abimael Guzmán, then a philosophy professor at the University of Huamanga in Ayacucho, lost an ideological dispute within the PCP over whether socialism should be pursued

through electoral politics or through armed revolution and broke away with twelve followers in Ayacucho and perhaps fifty elsewhere (CVR, 2003, p. 17). For the next decade, Guzmán’s faction operated as a clandestine political cell rooted in the Ayacucho university system, recruiting students who would take teaching posts in rural secondary schools and carry the movement’s doctrine into the countryside (CVR, 2003, pp. 17–19).

On May 17, 1982, five hooded Shining Path cadres broke into a polling station in the village of Chuschi, on the eve of Peru’s first democratic election after twelve years of military rule, and burned the ballot boxes. The act announced the beginning of what would become one of the deadliest civil conflicts in twentieth-century Latin America. By the time Guzmán was captured in September 1992, roughly 70,000 Peruvians had been killed, hundreds of thousands displaced, and one in three of the country’s districts touched by violence. Seventy-five percent of the dead were Indigenous, according to the TRC’s Informe Final (Laplante and Theidon, 2007), though Indigenous Peruvians constituted only about thirty percent of the population.

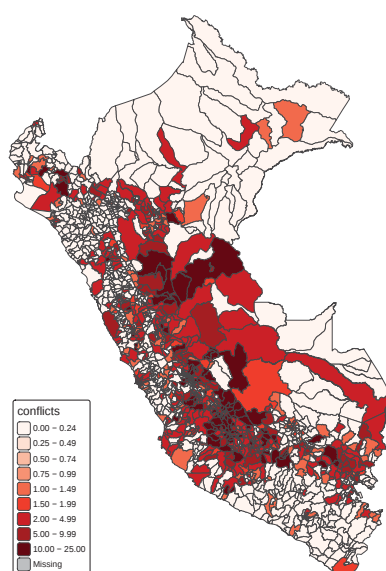


Figure 1. Violent events: 1980–2000

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

Three features of the conflict matter for what follows. The first is its geography. Figure 1 maps violent events at the district level. Violence was concentrated in the rural Andean highlands, particularly in the departments of Ayacucho, Huánuco, Junín, Apurímac, and Huancavelica, which are precisely the regions with the highest Indigenous population shares. This is the identification problem that motivates our border-sample design in Section 3: raw comparisons between violent

and non-violent districts would confound exposure to violence with pre-existing ethnic composition.

The second feature is the time path. Violence spiked sharply after 1983, when the Shining Path shifted from propaganda to a coercive campaign to impose control over the peasantry, killing community leaders, catechists, and peasants deemed insufficiently revolutionary (CVR, 2003, pp. 43–46). Counter-insurgency by state forces and locally-organized peasant patrols (*rondas campesinas*) intensified the toll. After a lull in the late 1980s, violence rose again until Guzmán’s capture in 1992 effectively ended the conflict. This compressed decade-long window of violence, bounded in calendar time but overlapping with the formative years of some cohorts and the post-formative years of others, is what makes the difference-in-differences strategy possible.

The third feature, and the one most central to our argument, is that the violence was overwhelmingly perpetrated by co-ethnics. The Shining Path recruited locally, often by force: each village was assigned a quota of young people, and families who refused were killed and their children enrolled anyway (CVR, 2003). Executions, kidnappings, and so-called “popular trials” were typically carried out by young cadres drawn from the same districts as their victims. The TRC’s *Informe Final* documents that the Senderistas operating in the Ayacucho highlands, where the insurgency began, were predominantly young men of peasant origin, recruited from the same rural communities in which they later carried out violence (CVR, 2003). Quechua peasants in Ayacucho faced violence not from an outgroup but from people who looked like them, spoke their language, and in many cases had grown up in the same hamlet. The 1983 massacre at Lucanamarca, in which a Shining Path column killed sixty-nine villagers with axes, machetes, and stones, was emblematic: survivors testified to the TRC that they recognized the perpetrators as fellow villagers, identifying them by their mothers’ surnames.² The TRC event-level data confirm this pattern quantitatively: 63 percent of Shining Path violence against Indigenous victims took place in the victim’s own birth district (Peru’s smallest administrative unit), and roughly 84 percent in the victim’s birth department.³ This intra-community character of the violence, documented more fully in Section 4, is what distinguishes the Peruvian case from the inter-ethnic conflicts studied in the existing literature, and drives our hypothesis about why Indigenous identity contracted rather than expanded under violence.

2.2. The Shining Path and Indigenous identity

The ideology that drove the Shining Path—*Pensamiento Gonzalo*—drew on Maoism filtered through the Peruvian socialist thinker José Carlos Mariátegui (Jima-González and Paradela-López, 2020, 2021). Mao’s signal contribution, from Guzmán’s perspective, was the identification of the peasantry rather than the industrial proletariat as the revolutionary subject. What the Chinese

² CVR, *Informe Final*, Tomo VII (La Masacre de Lucanamarca). The introduction quotes one such testimony.

³ Computed from the TRC actor-level database (*actos_est.sav*). See Section 4.1 and Appendix D for details.

Revolution had done for the category “peasant,” the Peruvian revolution would do again: the rural poor would be the engine of social transformation, and the countryside, not the industrial city, would be its base (Jima-González and Paradela-López, 2020).

This doctrinal choice had a specific implication for ethnic identity. If the peasant was the revolutionary subject, then every competing category of self-understanding, including ethnic, linguistic, regional, religious, was a form of false consciousness, an obstacle to class solidarity. The Shining Path’s 1988 *Fundamental Documents* (Shining Path, 1988) contain no reference to Indigenous culture, language, or territorial rights. The TRC concluded that the movement’s leadership, composed almost entirely of urban academics, regarded the Indigenous peasantry as a “homogeneous subject” and never engaged with its cultural specificity (CVR, 2003, p. 129). In a 1976 party document, the Shining Path declared that “the war belongs to the peasants under the leadership of the proletariat. . . . This is a peasant war, or it is nothing” (PCP, 1976). The category of Indigenous people was to be dissolved into the category of the peasant.

Violence was the instrument of that dissolution. *Pensamiento Gonzalo* held that the “quota of blood” was not an unfortunate cost of revolution but a necessary part of it (Jima-González and Paradela-López, 2020). In a 1974 speech to teachers, Guzmán told his audience that “war is necessary for the realization of the revolution” (Gúzman, 1974). A mid-level cadre put the logic to an interviewer more directly: rural populations who refused reeducation would receive “dictatorship, prison . . . death. The genocide which we have accepted, they will have to accept” (Wheat, 1990, p. 48). The massacres that followed, including the killing of the Uchuraccay leader Alejandro Huamán in 1982, of roughly eighty villagers in Lucanamarca in 1983, of fifty-five Asháninka peasants in 1993, were not departures from the movement’s doctrine but applications of it.

For the Indigenous population, this doctrine translated into a concrete choice. Abandon Indigenous practice, language, and self-identification and embrace the category of the peasant, or risk violence. Our empirical analysis asks whether this choice left a measurable imprint on ethnic identity a generation later.

2.3. When ethnic identity forms

Our identification strategy exploits the fact that ethnic identity is not formed at a single moment but over a developmental window. A large literature in developmental psychology places this window between early childhood and late adolescence. Ethnic awareness emerges around age four (Brand et al., 1974); ethnic self-categorization becomes stable in middle childhood (Davey and Mullin, 1980; Marks et al., 2007); active exploration and commitment to an ethnic identity are largely completed by the early twenties (Phinney, 1989, 1990; Yip, 2014). By the standard three-stage framework of Erikson (1968) and Marcia et al. (1980), an individual above age nineteen

or twenty has typically passed through unexamined identification, exploration, and commitment. After this age, ethnic self-concept becomes comparatively stable, responding less to environmental shocks.

This developmental window generates our empirical strategy. If violence reshapes ethnic identity, it should do so differentially across cohorts defined by age at exposure: individuals exposed while identity was still forming should diverge from individuals exposed only after identity had crystallized. The Shining Path conflict, by virtue of its compressed decade-long timing and spatially uneven reach, provides exactly this variation across individuals born in the same district.

The direction of the effect, however, is not obvious *ex ante*. In the leading account of conflict and ethnicity, developed theoretically by Rohner et al. (2013b) and empirically by Rohner et al. (2013a) for the Ugandan civil war, violence erodes generalized trust and drives individuals to retreat into their ethnic group, hardening identity and reproducing the conditions for future conflict. This account rests on a specific implicit assumption: that violence is inter-group, so that one's co-ethnics serve as a protective in-group against an external threat. The Peruvian case violates this assumption. Because Shining Path violence was primarily perpetrated by co-ethnics, often young men from the same villages as their victims, maintaining a visible Indigenous identity did not coordinate collective defense against an outsider. If anything, it marked individuals as members of precisely the community from which violence emanated. This reasoning predicts the opposite of the Rohner et al. (2013a,b) finding: Indigenous identity should contract, not expand, under intra-group violence. Testing this prediction is the task of the sections that follow.

3. Empirical Framework

3.1. Data

Our analysis combines two data sources: individual-level data on ethnic identity from Peru's National Household Survey (ENAHO), and georeferenced event-level data on acts of violence from the Peruvian Truth and Reconciliation Commission (TRC).

Ethnic identity. We measure Indigenous identity in two complementary ways, both drawn from the 2013–2017 waves of ENAHO, the first survey waves to include identity and language questions administered consistently across all 24 Peruvian departments. The first measure is self-identification: respondents are asked, with reference to their ancestry and customs, whether they consider themselves Quechua, Aymara, a member of another Indigenous community, Black, White, or Mestizo. We code an individual as Indigenous if the response falls in the first three categories and as non-Indigenous if the response is Mestizo. We exclude the small residual of Black and White respondents because the relevant margin of identity change in Peru is historically between Indigenous and Mestizo: *mestizaje*, the process of claiming mixed Spanish-Indigenous ancestry,

has long been the principal channel through which individuals shed Indigenous status.

The second measure is mother tongue: respondents are asked what language they first learned in childhood. We code an individual as Indigenous if the response is Quechua, Aymara, or another Indigenous language, and as non-Indigenous if the response is Spanish. Mother tongue is a behavioral rather than declarative measure—it captures a choice made by parents about what to transmit—and is therefore a useful complement to self-identification, which can shift with social incentives over the lifecycle. Figure A.1 in Appendix A shows the cohort-level share of each identity category from 1950 to 1992. Both measures trend toward less Indigenous identification in cohorts born during or after the conflict, a pattern consistent with our hypothesis and motivating the formal analysis that follows. Further detail on variable construction and the ENAHO modules used is provided in Appendix C.

Exposure to violence. The TRC database records 36,019 unique acts of violence, including killings, disappearances, illegal detentions, kidnappings, forced recruitments, torture, rapes, and woundings, drawn from 16,917 testimonies collected between the conflict’s beginning in 1980 and its effective end in 2000.⁴ Each event carries information on location (district), year, type of violence, and, in most cases, the ethnicity of the victim.

We collapse the TRC data to a district-year panel of violent-event counts across all 1,873 Peruvian districts and merge it to ENAHO using each individual’s district and year of birth. This produces our main regressor, $Exposure_i$, an indicator equal to one if individual i ’s district of birth experienced at least one violent event during 1980–1992, and zero otherwise. As reported by the TRC’s *Informe Final*, approximately seventy-five percent of fatal victims of the conflict spoke Quechua, Aymara, or another Indigenous language as their first language (Laplante and Theidon, 2007). Because our interest is in how violence shapes Indigenous identity specifically, we accordingly restrict attention to events in which the victim was Indigenous. In robustness checks (Section 3.5) we show that exposure to violence with non-Indigenous victims has no detectable effect on Indigenous identity, supporting the specificity of the mechanism.

Two measurement issues deserve explicit attention. First, because ENAHO records district of birth but individuals may have migrated, our exposure measure misclassifies those who left violent districts before violence occurred (false positives) and those who moved into violent districts after birth (false negatives). This attenuates our estimates toward zero. We address migration directly in Section 3.5.2, showing that (i) the main effects are concentrated among non-migrants and (ii) exposure to violence does not predict migration in our sample. Second, the TRC records events but

⁴ The 36,019 figure represents unique violent acts recorded in the full TRC database, spanning all perpetrators (Sendero Luminoso, state forces, rondas, and others) and all victim types, Indigenous and non-Indigenous alike. Because the TRC is based on individual testimonies, a single respondent may report multiple acts, for example, a witness describing several killings or a victim reporting both detention and torture, so the number of recorded acts exceeds the number of testimonies.

not all of them: the 36,019 figure is a lower bound on actual violence, reflecting the testimonies the Commission was able to collect. To the extent this underreporting is correlated with the dependent variable, it would also attenuate our estimates. Summary statistics for all variables are reported in Table A.1 in Appendix A.

3.2. Identification

The threat to identification is that Shining Path violence was not randomly distributed across districts. Figure 1 shows that violence concentrated in the rural Andean highlands—precisely the regions with the highest Indigenous population shares, the highest poverty rates, and the weakest state presence. A naive comparison of Indigenous identity between violent and non-violent districts would confound the effect of violence with all the pre-conflict differences that made some districts more attractive targets for the insurgency.

We address this in two steps. First, we restrict the sample to pairs of bordering districts that were similar on pre-conflict observables but experienced different levels of violence. Second, within this sample, we exploit variation in the *timing* of violence relative to the developmental window of identity formation, comparing cohorts exposed during formative years (ages 0–19) to cohorts exposed only after identity was largely formed (ages 20 and above).

The border-sample design follows Albertus (2020), whose study of land reform and conflict in Peru provides the geographic basis for our comparison. Peru’s 1969–1975 land reform was administered through thirteen “agrarian reform zones,” delineated by the state’s agrarian development agency (SIPA) on the basis of ecological conditions, transportation access, and existing market integration. Within each zone, districts were classified as “core” or “periphery” according to whether they shared a department with the zone’s regional administrative office, an assignment driven by administrative geography rather than by characteristics likely to predict later exposure to the Shining Path. Albertus shows that core districts received substantially more land redistribution than periphery districts, and that this difference in land reform generated a difference in subsequent conflict intensity: core districts, where grievance was partially addressed, saw fewer Shining Path events.

Our border sample consists of districts lying along the core–periphery boundaries *within* each agrarian reform zone. This design is close in spirit to Lichter et al. (2021), who use East German district borders to identify long-run effects of government surveillance. Two districts separated by a core–periphery boundary share the same agrarian zone and were drawn from the same pool of ecological and market conditions, but one received more land redistribution than the other and, consequently, experienced different violence intensity. Our identifying assumption is that, conditional on within-zone geography, this boundary is orthogonal to pre-conflict determinants of Indigenous identity.

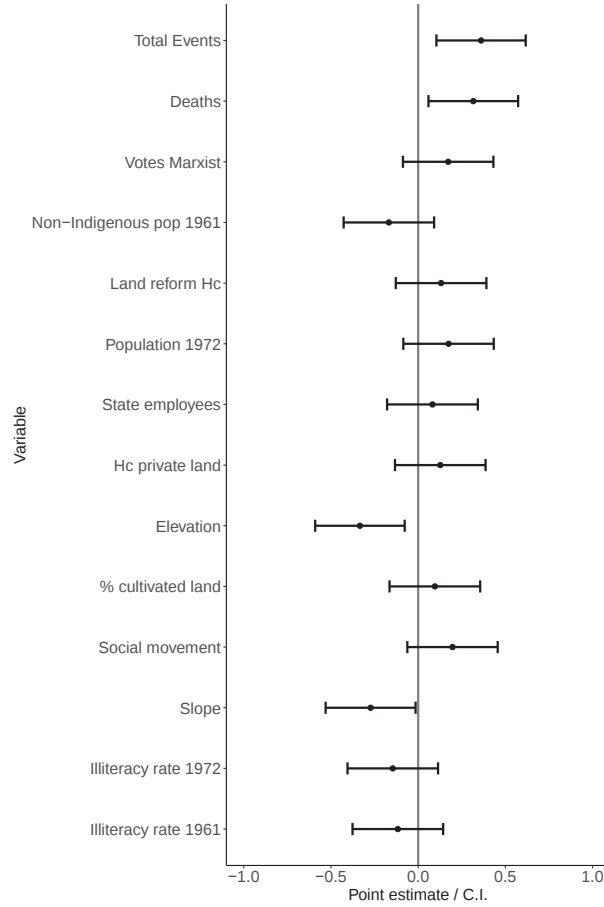


Figure 2. Pre-conflict balance: core vs. periphery districts in the border sample

Notes: The figure plots standardized mean differences between core and periphery districts in the border sample, with 95% confidence intervals. Base category is core districts. Unit of observation is the district. Variables are: total violent events, total deaths, share of votes to Marxist parties in 1980, share of non-Indigenous population in 1961, illiteracy rates in 1961 and 1972, number of pre-conflict peasant movements, average slope, share of cultivated land, hectares of private land, average elevation, 1961 state employees, 1972 population (thousands), and hectares of land reform. Statistical significance at 99%, 95%, and 90% confidence levels denoted: ***, **, *.

Figure 2 supports this assumption. We compute standardized mean differences between core and periphery districts in our border sample across fourteen pre-conflict characteristics. Core districts experience significantly more violent events and deaths, as expected. But they are statistically indistinguishable from periphery districts in the 1980 vote share for Marxist parties, a measure of pre existing political radicalism that might plausibly predict Shining Path recruitment, and, crucially, in the share of non Indigenous population in 1961. They are also balanced on 1961 and 1972 illiteracy rates, the number of pre-conflict peasant movements, 1972 population, share of cultivated land, hectares of private land, state employment, and total land reform hectares. Core districts do differ from periphery districts in average elevation and slope, but these are absorbed by district-of-birth fixed effects in our main specification.

A detailed discussion of how our border-sample design differs from Albertus's regression-discontinuity design is provided in Appendix E. We exploit *all* agrarian reform zones including

those he excludes, and we use the boundary as a source of sample restriction rather than as a running variable.

3.3. Empirical Strategy

Within the border sample, we exploit variation in the timing of violence relative to the developmental window of identity formation. The identifying idea is that individuals born in a violent district but past the identity-formation window by the time violence arrived should not experience effects of the same kind or magnitude as individuals born in the same district but exposed during their formative years. The comparison with individuals from non-violent bordering districts removes any systematic cohort trend in Indigenous identification that is common to the matched districts.

Our baseline specification is a difference-in-differences regression:

$$Y_{ikdtp} = \beta \text{Exposure}_p \times \text{Formative}_{ip} + \gamma_k + \gamma_t + \gamma_d + \gamma_p + \mathbf{x}_i' \delta + \epsilon_{ikdtp} \quad (1)$$

where Y_{ikdtp} is a binary indicator for the Indigenous identity of individual i , born in year k in district p , residing in district d , and surveyed in year t . We estimate the specification separately for our two measures: self-identification as Indigenous (versus Mestizo), and Indigenous mother tongue (versus Spanish).

Exposure_p equals one if district p experienced at least one Indigenous-victim violent event during 1980–1992, and zero otherwise. Formative_{ip} equals one if individual i was between the ages of 0 and 19 at the time a violent event occurred in district p , and zero if the individual was aged 20 or above (our baseline age category). The coefficient of interest, β , is the differential effect on Indigenous identity of being exposed to violence during the formative window, relative to exposure after identity has crystallized.

The fixed-effects structure absorbs several confounds. γ_k (year-of-birth) controls for any secular trend in Indigenous identification across cohorts which, as Figure A.1 shows, runs in the direction of less Indigenous identification over time. γ_t (year-of-interview) absorbs survey-wave effects. γ_p (district-of-birth) absorbs time-invariant district characteristics, including any pre-conflict determinant of Indigenous identity that is constant within a district. γ_d (district-of-interview) absorbs any effect of the current residential environment, which may differ from the birth district due to migration. $\mathbf{x}_i$ is a vector of individual-level controls including age, age squared, and sex. Given the birth-district fixed effects, identification of β comes from variation across cohorts within the same birth district, specifically from the interaction between whether the district was ever violent and whether a given cohort was in its formative window when the violence occurred. Standard errors are clustered at the district-of-birth by year-of-birth level, the level at

which treatment assignment varies.

The identifying assumption is that, absent the Shining Path violence, Indigenous identity in treated cohorts (those exposed during ages 0–19 in violent districts) would have followed the same trend as in the comparison cohorts (those exposed after age 19, and those in matched non-violent bordering districts). We test this assumption below using placebo cohorts.

Cohort-level specification. To characterize how the effect varies with age at exposure, and to test the parallel-trends assumption, we estimate a more flexible specification that replaces the binary $Formative_{ip}$ indicator with a set of cohort-specific interactions:

$$Y_{ikdtp} = \sum_{c \neq 20-24} \beta_c Exposure_p \times D_{ic} + \gamma_k + \gamma_t + \gamma_d + \gamma_p + \mathbf{x}'_i \delta + \epsilon_{ikdtp} \quad (2)$$

where D_{ic} is an indicator for whether individual i belonged to five-year cohort $c \in \{\text{post}, 0-4, 5-9, 10-14, 15-19, 25-29, 30-34\}$ at the time violence occurred in their birth district. The omitted category is the 20–24 cohort: individuals who were in their early twenties when violence arrived, just past the developmental window during which ethnic identity is typically formed. The “post” cohort captures individuals born in district p *after* the last recorded violent event there, a group whose identity formation postdates the conflict entirely.⁵

The coefficients β_c trace out the age profile of the effect of violence on Indigenous identity. Under our hypothesis, the effect should be large and negative for cohorts exposed during the formative window (0–4 through 15–19), absent for cohorts exposed after identity had crystallized (25–29 and 30–34), and plausibly absent for the post cohort whose identity formation occurred in a post-conflict environment. The post, 25–29, and 30–34 coefficients serve as placebos: a finding of no effect for these cohorts supports the parallel-trends assumption underlying our baseline estimates.

3.4. Main Results

Table 1 presents the baseline difference-in-differences estimates from Equation 1. Columns (1) and (3) use the full sample of Peruvian districts; columns (2) and (4) restrict to the border sample. Across both dependent variables, estimates in the border sample are larger in magnitude and more precisely identified than in the full sample, consistent with the border design removing the confounding of violence with pre-conflict Indigenous population shares.

In our preferred specification (columns 2 and 4), individuals exposed to Shining Path violence during their formative years are 3.4 percentage points less likely to speak an Indigenous language as their mother tongue, and 2.3 percentage points less likely to self-identify as Indigenous, than comparable individuals from the same birth districts whose identity was already formed at the time of exposure. Relative to sample means of 0.25 and 0.40, these represent reductions of 13.6 percent

⁵ For example, an individual born in 1987 in a district whose last recorded violent event was in 1984.

and 5.8 percent, respectively. Both estimates are statistically significant at conventional levels; the standard errors are clustered at the district-year level. Results are robust to clustering at the district-of-birth level (Table A.2) and to excluding the post-conflict cohort (Table A.3).

Table 1. Effect of Shining Path violence on Indigenous identity

	Indigenous language as mother tongue		Identify as Indigenous	
	all districts	border sample	all districts	border sample
	(1)	(2)	(3)	(4)
<i>Exposure × Formative</i>	−0.014** (0.006)	−0.034** (0.016)	−0.005 (0.005)	−0.023** (0.009)
District-of-interview FE	✓	✓	✓	✓
District-of-birth FE	✓	✓	✓	✓
Year-of-interview FE	✓	✓	✓	✓
Year-of-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean of dep. var.	0.23	0.25	0.36	0.40
SE cluster	district-year	district-year	district-year	district-year
<i>N</i>	208,768	33,326	168,301	27,842

Notes: The table reports estimates from Equation 1. The unit of observation is the individual. The dependent variable in columns (1)–(2) is an indicator for whether the individual speaks an Indigenous language as their mother tongue; in columns (3)–(4) it is an indicator for self-identification as Indigenous. Columns (1) and (3) use all Peruvian districts; columns (2) and (4) restrict to the border sample described in Section 3.2. All specifications include year-of-birth, year-of-interview, district-of-interview, and district-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors are clustered at the district-year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Cohort profile and parallel trends. Figure 3 presents the cohort-level estimates from Equation 2, with the 20–24 cohort as the omitted baseline. Two patterns are visible. First, the negative effect of violence on Indigenous identity is concentrated primarily in early childhood, especially for mother tongue transmission. Individuals exposed between ages 0 and 4 are 7.8 percentage points less likely to speak an Indigenous mother tongue and 4.0 percentage points less likely to self-identify as Indigenous than comparable individuals from the baseline cohort. For mother tongue, the effect attenuates at ages 5–9 (4.4 percentage points) and becomes statistically indistinguishable from zero by ages 10–14 and 15–19. For self-identification, the pattern is weaker and less precisely estimated: the largest decline also appears in the 0–4 cohort, but later childhood and adolescent coefficients are small and statistically indistinguishable from zero throughout.⁶ This age gradient is strongest for mother tongue, the outcome most directly shaped by parental transmission, and is consistent with the developmental psychology evidence summarized in Section 2: early childhood is the period when language acquisition is most malleable, while broader self-identification appears comparatively less responsive and more imprecisely measured in these cohort specifications.

⁶ These magnitudes come from Table A.6 in the appendix, which reports the full set of coefficients.

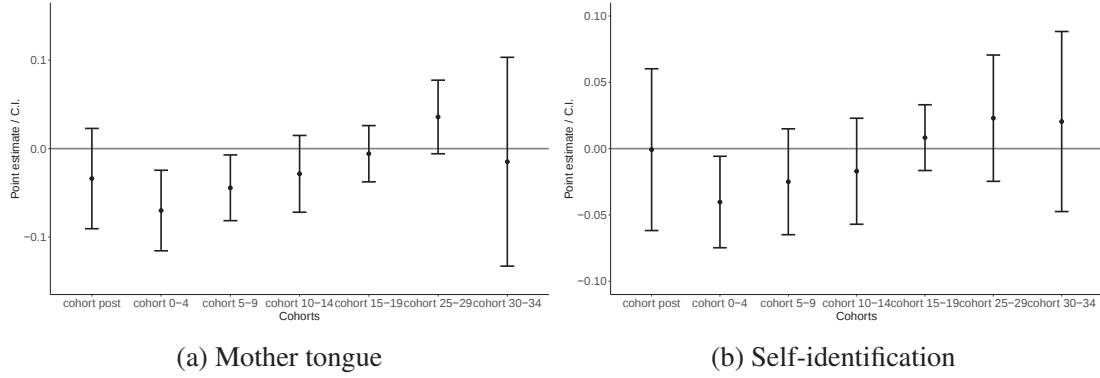


Figure 3. Effect of violence on Indigenous identity, by age at exposure

Notes: The figure plots β_c coefficients from Equation 2 with 95 percent confidence intervals. The omitted cohort is 20–24. The “post” cohort consists of individuals born in district p after the last recorded violent event in that district. Panel (a) uses Indigenous mother tongue as the dependent variable; panel (b) uses self-identification as Indigenous. The sample is restricted to the border sample. All specifications include year-of-birth, year-of-interview, district-of-interview, and district-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors are clustered at the district-year level.

Second, the placebo cohorts behave as the identifying assumption requires. The estimates for cohorts 25–29 and 30–34, corresponding to individuals whose core ethnic identity was already formed when violence arrived, are small and statistically indistinguishable from zero for self-identification, while mother tongue effects are likewise largely attenuated relative to early childhood exposure. The coefficient on the post cohort, that is, individuals born in violent districts after conflict had ended, is also statistically indistinguishable from zero in the border sample specifications. Together, these placebo patterns support the interpretation that the main estimates capture the disruptive effect of violence on identity formation rather than broader secular divergence between violent and non-violent districts.

The finding that exposure to violence during the first years of life produces the largest effect, while exposure during adolescence does not, is the headline result of the paper. It tells us that Shining Path violence reshaped ethnic identity, and that it did so through a channel that operated primarily during the years in which language and self-concept are first transmitted. We turn in Section 4 to the question of which channel this is and why it should produce a *decline* in Indigenous identity rather than a strengthening.

3.5. Robustness

We consider four threats to our interpretation. First, the estimated effect could reflect a generic response to violence rather than something specific to violence against Indigenous victims. Second, migration could confound the estimates if people who left violent districts differ systematically from those who stayed. Third, the effect might be part of a broader contraction of all identities, rather than a specific decline in Indigenous identity. Fourth, the effect could be concentrated in particular

Indigenous subgroups in ways that bear on the mechanism. We address each in turn.

3.5.1. Violence against non-Indigenous victims

Our main specification restricts attention to violent events in which the victim was Indigenous. If the decline in Indigenous identity we estimate reflects a generic response to exposure to violence, such as anxiety, displacement, disruption of social networks, then violence against non-Indigenous victims should produce similar effects. If, instead, the effect operates through the specific dynamics of violence directed at Indigenous communities, violence against non-Indigenous victims should have no effect.

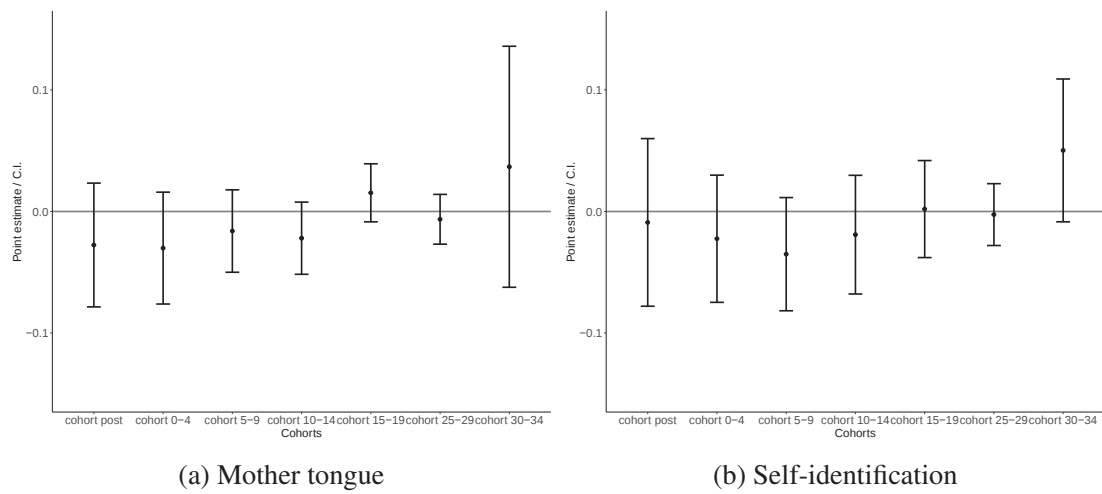


Figure 4. Placebo: exposure to violence against non-Indigenous victims

Notes: The figure plots β_c coefficients from Equation 2 estimated using exposure to violence against non-Indigenous victims as the treatment, with 95 percent confidence intervals. The omitted cohort is 20–24. Panel (a) uses Indigenous mother tongue as the dependent variable; panel (b) uses self-identification as Indigenous. Sample restricted to the border sample. All specifications include year-of-birth, year-of-interview, district-of-interview, and district-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors clustered at the district-year level.

Figure 4 replicates the cohort-level specification of Equation 2, replacing Indigenous-victim exposure with non-Indigenous-victim exposure. Across all age cohorts and both dependent variables, the coefficients are small, statistically indistinguishable from zero, and show no age profile. Exposure to non-Indigenous violence has no detectable effect on Indigenous identity. The contrast with our main estimates (Figure 3) is sharp: whatever channel drives the decline in Indigenous identity, it operates specifically when violence is directed at Indigenous victims.

3.5.2. Migration

Conflict displaces people, and displacement could confound our estimates.⁷ In our sample, individuals exposed to violence in their birth district are 8.4 percentage points more likely to have moved away by the time of the ENAHO survey (Table A.4, column 7). If migrants shed Indigenous identity upon moving to urban areas, where Indigenous identity is socially less salient, our main estimates could reflect selection into migration rather than a direct effect of violence.

We address this in two steps. First, we re-estimate our baseline specification controlling for a $Migration_i$ indicator equal to one if the individual currently resides in a different district from their birth district. Second, we interact $Migration_i$ with $Exposure_p \times Formative_{ip}$ to allow the effect of violence to differ for migrants and non-migrants. Third, we test whether exposure to violence itself predicts migration within the border sample.

Table A.4 reports the results. The main coefficient is essentially unchanged when $Migration_i$ is included as a control (columns 2 and 5): individuals exposed to violence during their formative years remain 3.4 percentage points less likely to speak an Indigenous mother tongue and 2.3 percentage points less likely to self-identify as Indigenous. The interaction specification (columns 3 and 6) shows that the effect is concentrated among non-migrants: individuals who were exposed to violence and remained in their birth districts are 5.1 percentage points less likely to speak an Indigenous mother tongue and 3.6 percentage points less likely to self-identify as Indigenous, relative to non-exposed non-migrants. The effect for exposed migrants, relative to the same baseline, is statistically indistinguishable from zero—consistent with an interpretation in which migrants had already left the environment in which the identity shift was most likely to occur. Finally, when we use $Migration_i$ as the outcome and estimate Equation 1 (column 8), the coefficient is a precise zero: within the border sample, formative-age exposure to violence does not predict subsequent migration. Selection into migration is therefore not driving our results.

3.5.3. Other identities

Our main finding is that violence reduces Indigenous identification. A sharper interpretation is that violence reduces identification with any category that the violence sought to impose or associate with Indigenous communities. To test which of these interpretations fits the data, we estimate analogous specifications using class identity and national identity as outcomes.

We measure class identity through membership in class-associated organizations, including workers' associations, professional associations, political parties, and peasant communities, using data from ENAHO. National identity is measured using the “How proud are you to be Peruvian?”

⁷ On the conflict-migration link, see Morrison (1993) on Guatemala, Basu and Pearlman (2017) on Mexico, Silva and Massey (2015) on Colombia, and De Vries and Guild (2019) on Africa.

question in the Latinobarómetro surveys. Table 2 reports the class-identity results. Formative-age exposure to violence has no detectable effect on membership in workers' associations, professional associations, or political parties. It does, however, reduce the probability of belonging to a peasant community by 2.3 percentage points, a striking result given that the Shining Path explicitly sought to impose peasant identity. Violence did not produce embrace of the imposed identity; it produced rejection of it.

Table 3 reports the national-identity results. Individuals exposed to violence in their formative years are 5.2 to 9.8 percentage points less likely to express pride in Peruvian identity, depending on the measure, with the effect robust to controlling for Indigenous status. Read jointly, these results describe a consistent pattern: violence during the formative years produces a retreat from publicly held collective identities, including Indigenous, peasant, national identities, rather than a specific rejection of the Indigenous category alone. The mechanism we identify in Section 4 helps explain why this retreat is most pronounced for Indigenous identity: because Indigenous identity was the most visible marker of membership in the community from which violence emanated.

Table 2. Effect of violence on class-identity measures

	Workers' association	Professionals' association	Political party	Peasant community
	(1)	(2)	(3)	(4)
<i>Exposure × Formative</i>	0.008 (0.008)	0.005 (0.009)	0.002 (0.005)	−0.023** (0.011)
District-of-interview FE	✓	✓	✓	✓
District-of-birth FE	✓	✓	✓	✓
Year-of-interview FE	✓	✓	✓	✓
Year-of-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean of dep. var.	0.037	0.37	0.011	0.22
<i>N</i>	13,868	13,868	13,868	13,868

Notes: The table reports estimates from Equation 1. The unit of observation is the individual. Dependent variables are indicators for membership in the specified organization, drawn from ENAHO. Sample restricted to the border sample. All specifications include year-of-birth, year-of-interview, district-of-interview, and district-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors clustered at the district-year level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

3.5.4. Heterogeneity across Indigenous groups

Peru is home to more than fifty Indigenous groups, but three categories dominate our data: Quechua (22 percent of the sample), Aymara (0.6 percent), and other Indigenous (2.9 percent).

Table 3. Effect of violence on Peruvian national identity

	Pride indicator		Pride, ordered	
	(1)	(2)	(3)	(4)
<i>Exposure × Formative</i>	−0.057*** (0.021)	−0.052** (0.020)	−0.099*** (0.036)	−0.098** (0.038)
Non-Indigenous		0.084*** (0.027)		0.177*** (0.044)
District FE	✓	✓	✓	✓
Survey-wave FE	✓	✓	✓	✓
Year-of-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean of dep. var.	0.83	0.83	3.38	3.38
<i>N</i>	2,675	2,656	2,675	2,656

Notes: The table reports estimates from Equation 1, with Peruvian national identity as the dependent variable. Columns (1)–(2) use a binary indicator equal to one if the respondent is “very proud” or “proud” to be Peruvian; columns (3)–(4) use the full ordered response (1 = not proud, 4 = very proud). Data are from the 2002–2015 waves of the Latinobarómetro. Columns (2) and (4) include an indicator for Non-Indigenous identity. All specifications include district, survey-wave, and year-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors clustered at the district-year level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimating our main specification separately by group reveals a clear pattern (Table A.5). Formative-age exposure to violence significantly reduces Quechua identification and identification with other smaller Indigenous groups, but has no detectable effect on Aymara identification.

The null result for Aymara is partly statistical, as fewer than 200 Aymara observations are available in the border sample, which leads to large standard errors. But there is also a substantive reading that links to the mechanism we develop in the next section. The Aymara are geographically concentrated in the southern Andes, in regions that were relatively distant from the core zones of Shining Path activity in Ayacucho, Huánuco, and Junín. Even within districts that experienced violence, Aymara individuals were more likely than Quechua individuals to face perpetrators who were not co-ethnics, because Quechua cadres and non-Indigenous cadres dominated the Shining Path’s regional operations. If the mechanism driving our main results is that visible Indigenous identity marks one as a member of the community from which violence comes, the marker only operates when victim and perpetrator share the identity. For Aymara victims of mostly non-Aymara perpetrators, the logic does not apply. We return to this argument in Section 4.

4. Mechanism

Our main results present an empirical puzzle. In the leading account of conflict and ethnicity, developed theoretically by Rohner et al. (2013b) and validated empirically by Rohner et al. (2013a) for the Ugandan civil war, violence erodes generalized trust, pushes individuals to retreat into their ethnic group, and hardens ethnic identification. This is the “war trap”: conflict and ethnic attachment reinforce each other, and societies fall into persistent, self-sustaining cycles of ethnic violence. Our findings for Peru run in the opposite direction. Exposure to violence during formative years *reduces* Indigenous identification rather than strengthening it. And, consistent with this, the Peruvian conflict did not reproduce itself: as Figure 5 shows, violence declined rapidly after Guzmán’s capture in 1992 and essentially ended by 1995. There is no war trap in Peru.

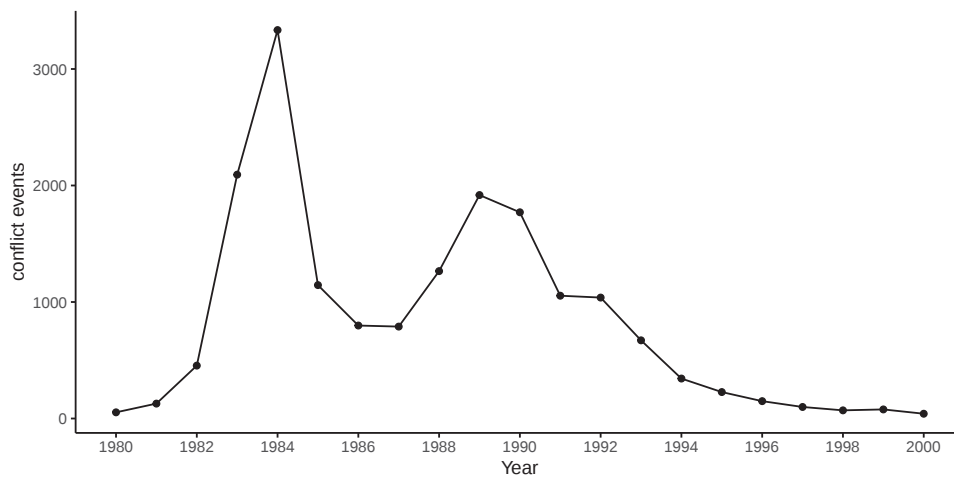


Figure 5. Violent events by year, 1980–2000

Notes: Number of violent events in Peru by year, 1980–2000, at the national level. Events include killings, disappearances, illegal detentions, kidnappings, forced recruitments, torture, rapes, and woundings. Data from Peru’s Truth and Reconciliation Commission.

The difference between Peru and the cases that motivate the Rohner et al. (2013a,b) framework lies in who the perpetrators were. In the Ugandan conflict, violence was inter-ethnic: rebels and government forces drew from different ethnic groups than their victims. Under such conditions, one’s co-ethnics are a natural in-group. They share the experience of violence and offer collective defense, and visible ethnic identity coordinates this defense. In the Peruvian case, by contrast, violence was overwhelmingly *intra-community*, a pattern we document below.

4.1. The local character of Shining Path violence

A distinctive feature of the Peruvian conflict, relative to most civil wars analyzed in the political-economy literature, is that perpetrators and victims typically came from the same place.

The Shining Path recruited cells locally, often by force, and deployed them to carry out violence in the communities from which they had been drawn. We document this pattern directly in the TRC event-level database.

The TRC records each violent act, the perpetrator group responsible, and the geographic location of the event. Crucially, for Indigenous victims it also records the victim's district of birth, allowing us to compare the geographic location of each Indigenous-victim event to the victim's own origin district. The TRC does not record the district of origin of perpetrators. But if Shining Path violence was, as the qualitative literature argues, carried out by cadres who operated in the districts where they lived, then violence against an Indigenous victim in the victim's own birth district is necessarily violence perpetrated on that victim's home ground, and is with high probability perpetrated by members of the victim's own community.

Among the 7,674 Shining Path victim-acts in the TRC event-level file with an Indigenous mother tongue and birthplace recorded, 63.6 percent took place in the victim's own birth district (Peru's smallest administrative unit), 74.8 percent in the victim's birth province, and 83.6 percent in the victim's birth department.⁸ These figures are conservative lower bounds: migration between birth and event location, by either the victim or the perpetrator before the conflict, would attenuate the observed match rate even if all perpetrators were truly co-local.

Under the assumption that Shining Path cadres were geographically local, roughly two thirds of the Sendero violence against Indigenous victims that the TRC records at usable geographic precision was carried out in the victim's own home district. The scale of the local-perpetrator pattern is not a secondary feature of the conflict; it is a defining one.

This observation suggests our mechanism. When violence is inter-group, visible ethnic identity buys protection; when violence is intra-group, visible ethnic identity buys nothing and may mark one as embedded in the community from which violence emanates. Parents facing violence from co-ethnic perpetrators have a reason to withhold the transmission of language and custom to young children, a margin that is theirs to control, unlike the ethnicity of the perpetrators. This predicts exactly what we find: a decline in Indigenous identification concentrated in early childhood, when parental transmission is most consequential, generated specifically by violence against Indigenous victims, and absent when the violence is against non-Indigenous targets.

In the remainder of this section, we test this mechanism in two steps. We first show that the pattern of trust responses in Peru differs from the Rohner et al. (2013a,b) benchmark in a way consistent with intra-group violence. We then provide direct evidence that the effect of violence on Indigenous identity is concentrated in Indigenous-homogeneous districts, where, under our

⁸ Computed from the TRC event-level database (*actos_est.sav*). Event-level figures, which collapse multiple victims within the same event into a single observation, are nearly identical and are reported in Appendix Table A.12. See Appendix D for full construction details.

mechanism, violence was most likely to be intra-group from the victim's perspective.

4.2. Trust

The Rohner et al. mechanism predicts that violence erodes generalized trust (trust in strangers outside one's community) and raises interpersonal trust (trust in neighbors and kin), driving individuals into the arms of their co-ethnics. Intra-group violence should produce the opposite pattern: erosion of interpersonal trust, since neighbors and kin are the source of threat, with generalized trust left comparatively intact.

We test these predictions using three data sources: the Latin American Public Opinion Project surveys (LAPOP, 2006–2016) for interpersonal trust, the Latinobarómetro surveys (2002–2015) for generalized trust, and ENAHO module 85 (2013–2017) for institutional trust. A methodological caveat applies to LAPOP and Latinobarómetro: both are representative at the province level (Peru's second administrative layer, with 196 units) rather than the district level (1,873 units), and neither records district or province of birth. We therefore assume residence equals birthplace and construct province-level exposure indicators. Because our main results show that the effect is concentrated in non-migrants and Indigenous-victim violence, both approximations should push coefficients toward zero. Estimates reported here are best interpreted as lower bounds on the true effects.

Interpersonal trust. Table A.8 in the appendix reports estimates using LAPOP's standard interpersonal trust item ("How trustworthy are the people in this community?") as the dependent variable.⁹ Individuals exposed to violence during their formative years report interpersonal trust 0.17 units lower than comparable individuals not exposed, a 6.5 percent reduction relative to the sample mean of 2.46. The effect persists when we condition on Non-Indigenous status (a partial control given the possibility that ethnic identification is itself an outcome of violence).

This contrasts sharply with Rohner et al. (2013a), who find that violence in Uganda raised interpersonal trust within villages. The opposite finding in Peru is consistent with intra-group violence: when the people nearest to you are the source of threat, local trust erodes rather than consolidating.

Generalized trust. Table A.9 in the appendix reports analogous estimates using the Latinobarómetro's generalized trust item ("Generally speaking, would you say that you can trust most people. . . ?"). The coefficient is small, positive, and statistically indistinguishable from zero (1.4 percentage points, s.e. similar magnitude). We cannot reject the null of no effect. This too contrasts with Rohner et al., who find that violence *decreases* generalized trust. In Peru, violence appears to have left generalized trust essentially intact while eroding interpersonal trust—again, the pattern predicted by intra-group violence rather than inter-group violence.

⁹ The original four-point response scale runs from 1 (very trustworthy) to 4 (untrustworthy); we reverse it for interpretability, so higher values indicate more trust.

Institutional trust. Finally, we use ENAHO module 85 to measure trust in the district government, national government, judiciary, congress, and Catholic Church. Table A.10 in the appendix reports null results across all five institutions. Violence does not detectably change institutional trust in our sample. This is useful because it bounds the trust effects: the erosion is specifically *interpersonal* (trust in people one encounters in daily life) rather than a generalized collapse of trust in all social structures. The specificity is consistent with our mechanism, in which the relevant trust erosion is in the local community where violence was experienced.

4.3. Within-group versus between-group violence

The trust evidence is suggestive but indirect. For a sharper test of the mechanism, we need a proxy for whether, from an Indigenous individual’s perspective, the violence in their district was more likely to be intra-group or inter-group. We construct such a proxy using Indigenous ethnic diversity at the district level.

Peru is home to more than fifty Indigenous groups, though Quechua, Aymara, and a residual category of other Indigenous peoples dominate the data. In districts where essentially one Indigenous group is present, any violence involving an Indigenous perpetrator and an Indigenous victim is necessarily intra-group. In districts with multiple Indigenous groups, violence can cross group boundaries even when both sides are Indigenous. The Shining Path was not an ethnic movement, so its cadres were not drawn from any particular group with consistency; what varies across districts is therefore not the perpetrator’s specific ethnicity but the probability that, from a given victim’s perspective, the local violence is perceived as coming from within or outside their specific Indigenous community.

Following Alesina et al. (1999), we construct a district-level Indigenous diversity index equal to the probability that two randomly drawn Indigenous individuals from the same district belong to different Indigenous groups. We compute this index in two ways. The first uses ENAHO respondents born before 1950, cohorts whose ethnic identification was established well before the Shining Path’s first armed action in 1982. The second uses the 1981 Peruvian census, the last full enumeration before the conflict began. The census-based measure has two virtues. It is mechanically immune to any post-conflict reshaping of which Indigenous groups remained in a district, since it is fixed in time before the insurgency. And it covers the full district population rather than survey respondents, eliminating the small-sample noise that affects the ENAHO-based index in sparsely surveyed districts.

Our mechanism makes a clear prediction: the negative effect of violence on Indigenous identity should be concentrated in Indigenous-homogeneous districts, where intra-group violence was most likely, and should attenuate in diverse districts, where Indigenous victims could interpret violence as coming from other Indigenous groups. We test this by interacting $Exposure_p \times Formative_{ip}$ with

the diversity measure. In this specification, the base category is individuals in fully homogeneous districts not exposed to violence.

Table 4 reports the results. Columns (1) and (2) use the ENAHO-based diversity index; columns (3) and (4) replace it with the 1981 census-based index, with no other change to the specification.¹⁰ The mother tongue results in columns (1) and (3) strongly support the prediction. In fully homogeneous districts, exposure during formative years reduces the probability of speaking an Indigenous mother tongue by 5.8 percentage points using the ENAHO measure and 5.4 percentage points using the census measure. The interaction with Indigenous diversity is large, positive, and significant at the one percent level in both specifications: +0.213 with the ENAHO measure and +0.133 with the census measure. A one-standard-deviation increase in Indigenous diversity (0.17) offsets this negative effect by approximately 3.6 percentage points under ENAHO and 2.3 percentage points under the census measure, implying that greater diversity substantially attenuates the assimilation effect of violence exposure. In sufficiently diverse districts, the estimated effect can even reverse sign entirely. These mother-tongue coefficients, which most directly capture parental language-transmission decisions, are therefore highly sensitive to local ethnic structure and the implied likelihood of intra-group versus inter-group threat.

Columns (2) and (4), with self-identification as the outcome, show a consistent but weaker pattern. The main effect is negative and statistically significant in both specifications (−0.029 and −0.024), but the interaction terms are small and statistically insignificant for both diversity measures (+0.042 with ENAHO; +0.015 with the census). As in earlier versions of this specification, this suggests that the diversity-mediated mechanism operates most clearly through language transmission rather than self-identification, where estimates remain too imprecise to detect a comparable moderating effect at current sample sizes.

That the qualitative pattern survives almost intact when the diversity index is recomputed from the 1981 census is reassuring on two counts. First, the mother-tongue interaction is not an artifact of how the ENAHO-based index is constructed: replacing the survey measure with a pre-conflict census measure preserves both the sign and the significance, with magnitudes that are smaller but in the same range. Second, and more substantively, it rules out an interpretation in which the diversity index is itself a post-conflict object, shaped by selective migration or mortality during the violence. The 1981 measure is fixed before the Shining Path's first armed action; the result it delivers is therefore a true cross-district comparison of pre-conflict diversity, not a reflection of how the conflict redistributed Indigenous groups across space.

¹⁰ The 1981 census files cover 136 of 229 birth districts in the border sample; the remaining 93 districts are dropped from the columns (3)–(4) sample. Unmatched districts are located in the conflict's core departments (Ayacucho, Apurímac, Huancavelica), which under-represents high-intensity exposure variation. Sample restrictions of this kind typically bias interaction coefficients toward zero, so the census-based estimates should be read as conservative relative to the full-coverage survey-based ones.

This is also a useful framing for the null result on Aymara that we noted in Section 3.5. The Aymara are geographically concentrated in the southern Andes, largely outside the Shining Path’s main operational zones in Ayacucho, Huánuco, and Junín. Aymara victims of Shining Path violence were, to a greater extent than Quechua victims, facing perpetrators who were not their co-ethnics. Our mechanism predicts that the identity-contraction effect should be muted in their case, which, with the caveat of limited statistical power, is what we find.

Table 4. The effect of violence on Indigenous identity, by Indigenous diversity of the birth district

	Mother tongue	Identify Indigenous	Mother tongue	Identify Indigenous
	(1)	(2)	(3)	(4)
<i>Exposure × Formative</i>	−0.058*** (0.019)	−0.029*** (0.011)	−0.054** (0.026)	−0.024* (0.014)
× Diversity index	0.213*** (0.057)	0.042 (0.053)		
× Diversity index (1981 census)			0.133*** (0.049)	0.015 (0.034)
District-of-interview FE	✓	✓	✓	✓
District-of-birth FE	✓	✓	✓	✓
Year-of-interview FE	✓	✓	✓	✓
Year-of-birth FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
SE cluster	district-year	district-year	district-year	district-year
<i>N</i>	29,911	25,566	22,055	18,840

Notes: The table reports estimates from Equation 1, interacting $Exposure_p \times Formative_{ip}$ with the continuous Alesina et al. Indigenous diversity index in the birth district. The index measures the probability that two randomly drawn Indigenous individuals from the district belong to different Indigenous groups. In columns (1) and (2), the index is constructed from ENAHO respondents born before 1950. In columns (3) and (4), the index is constructed from the 1981 Peruvian census, the last full enumeration before the conflict began. Sample restricted to the border sample. The smaller sample in columns (3) and (4) reflects districts in the border sample that could not be matched to a 1981 census district. All specifications include year-of-birth, year-of-interview, district-of-interview, and district-of-birth fixed effects. Controls include age, age squared, and sex. Standard errors clustered at the district-year level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.4. Synthesis

Four pieces of evidence converge on the intra-group-violence mechanism. First, the violence in question was overwhelmingly local: 63 percent of Shining Path violence against Indigenous victims occurred in the victim’s own birth district, establishing the empirical premise that perpetrators and victims came from the same communities (Section 4.1). Second, exposure to non-Indigenous violence has no effect on Indigenous identity (Section 3.5): the mechanism is triggered specifically by violence against Indigenous victims, not by violence in general. Third, the trust response in Peru is the reverse of the Rohner et al. (2013a,b) pattern: interpersonal trust falls while generalized

trust does not, consistent with violence coming from within the local community. Fourth, and most directly, the identity effect on mother tongue is concentrated in Indigenous-homogeneous districts, where intra-group violence was most likely, and attenuates in diverse districts where it was not. The interaction is robust to constructing the diversity index from the 1981 census, fixed before the conflict began, ruling out an interpretation in which the diversity measure is itself an artifact of the violence. Read together, the results suggest that ethnic identity is durable when the threat comes from outside the group, but fragile when it comes from within.

5. Conclusion

This paper shows that a decade of civil violence can produce a measurable decline in the ethnic identity of exposed cohorts three decades later. Using the Shining Path insurgency in Peru as a natural experiment, we find that individuals exposed to Indigenous-victim violence during the formative years of identity development are substantially less likely to speak an Indigenous mother tongue or to self-identify as Indigenous than cohorts whose identity was already formed when violence arrived. The effect is concentrated in early childhood, survives a range of robustness checks including placebos for non-Indigenous violence, migration, and alternative identities, and is absent for cohorts unexposed during the developmental window.

The mechanism is not a generic trauma response but something more specific. The Shining Path recruited locally and coerced violently; in the villages where violence was heaviest, perpetrators and victims came from the same ethnic community. Under these conditions, visible Indigenous identity did not coordinate collective defense against an external enemy, the classic channel through which violence is thought to harden ethnic cohesion, but marked individuals as embedded in the community from which violence emanated. Consistent with this, we find that the effect is concentrated in Indigenous-homogeneous districts where intra-group violence was most likely, that interpersonal trust (but not generalized trust) fell in response to exposure, and that violence did not produce embrace of the class identity the Shining Path sought to impose. The evidence describes a specific mechanism: when violence comes from within the group, parents respond by withholding the cultural markers that would otherwise be transmitted to children.

This result overturns a widely held prediction about how conflict reshapes ethnic identity. The leading account, developed by Rohner et al. (2013b) and validated in inter-ethnic settings by Rohner et al. (2013a), holds that violence hardens ethnic attachment, pushing societies into self-sustaining cycles of conflict, a “war trap.” Our findings suggest that this result is not general. Ethnic identity is durable when the threat comes from outside the group, but fragile when it comes from within. The Peruvian conflict’s failure to reproduce itself across generations, as violence fell sharply after Guzmán’s capture and did not return, is consistent with this mechanism: where violence erodes

rather than consolidates ethnic cohesion, the conditions for persistent conflict are not reproduced.

More broadly, the paper speaks to the economics of identity. A large literature following Akerlof and Kranton (2000) has established that identity shapes economic behavior, but the economic determinants of ethnic identity itself have received less attention. Treating ethnicity as a fixed trait, as much of the empirical literature implicitly does, misses the margins along which individuals and families respond to incentives. We show that a decade of targeted violence is enough to shift those margins decisively, and that the direction of the shift depends on who is doing the violence. Where ethnic categorization is contested, identity is also endogenous to the security of the category itself.

These findings have implications beyond Peru. In conflicts ranging from Rwanda to Myanmar to the former Yugoslavia, the ethnic identity of perpetrators relative to their victims has varied, and our results suggest this variation matters for the long-run cultural trajectory of the societies involved. They also raise a question our data cannot answer: whether the identity shifts we document are stable across generations, or whether children born to parents who withheld Indigenous transmission can recover the identity their grandparents held. The answer matters for how societies recover from civil conflicts in which identity was itself a target.

References

- Akerlof, George A and Rachel E Kranton (2000). Economics and identity. *The Quarterly Journal of Economics* 115(3), pp. 715–753.
- (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.
- 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*.

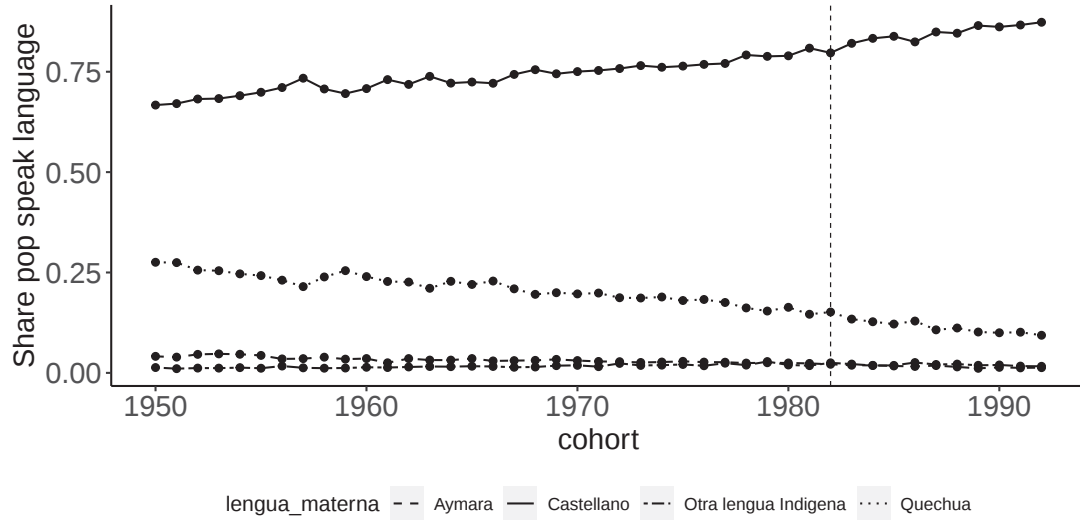
- Brand, Elaine S, Rene A Ruiz, and Amado M Padilla (1974). Ethnic identification and preference: A review. *Psychological Bulletin* 81(11), p. 860.
- Chen, Yan and Sherry Xin Li (2009). Group identity and social preferences. *American Economic Review* 99(1), pp. 431–57.
- Clots-Figueras, Irma and Paolo Masella (2013). Education, language and identity. *The Economic Journal* 123(570), F332–F357.
- 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, Abigail (1974). *LA PROBLEMÁTICA NACIONAL*. URL: https://cedema.org/digital_items/631 (visited on 05/12/2022).
- Huntington, Samuel P (2006). *Political order in changing societies*. Yale university press.
- Jima-González, Alexandra and Miguel Paradelo-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.
- Kranton, Rachel E (2016). Identity economics 2016: Where do social distinctions and norms come from? *American Economic Review* 106(5), pp. 405–09.
- 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 Sieglöcher (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.
- 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, Mathias Thoenig, and Fabrizio Zilibotti (2013a). Seeds of distrust: Conflict in Uganda. *Journal of Economic Growth* 18(3), pp. 217–252.
- (2013b). War signals: A theory of trade, trust, and conflict. *Review of Economic Studies* 80(3), pp. 1114–1147.
- Shining Path (1988). *Documentos fundamentales [Fundamental Documents]*. URL: http://www.solrojo.org/pcp_doc/pcp_gd88.htm (visited on 05/12/2022).

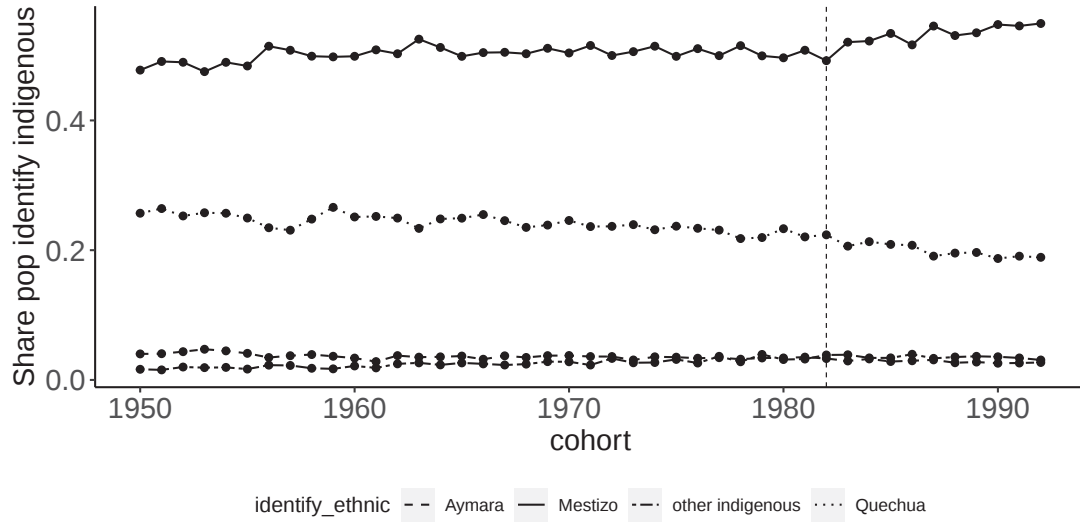
- Silva, Adriana Carolina and Douglas S Massey (2015). Violence, networks, and international migration from Colombia. *International Migration* 53(5), pp. 162–178.
- Wheat, Andrew (1990). Shining Path's "fourth sword" ideology. *Journal of Political & Military Sociology*, pp. 41–55.
- Wood, Elisabeth Jean (2003). Insurgent collective action and civil war in El Salvador. Cambridge University Press.
- Yip, Tiffany (2014). Ethnic identity in everyday life: The influence of identity development status. *Child Development* 85(1), pp. 205–219.

Appendix

A. Tables and Figures



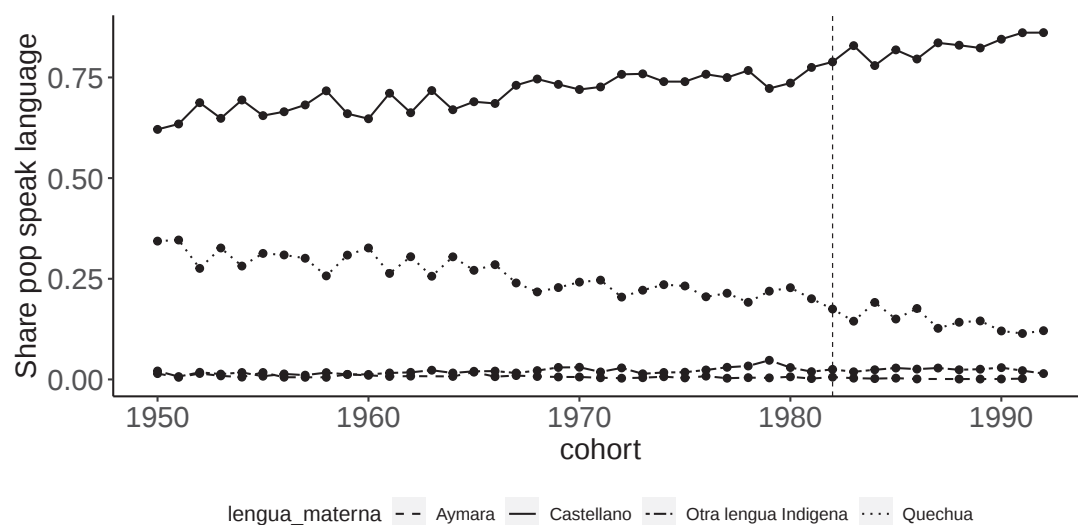
(a) Indigenous language



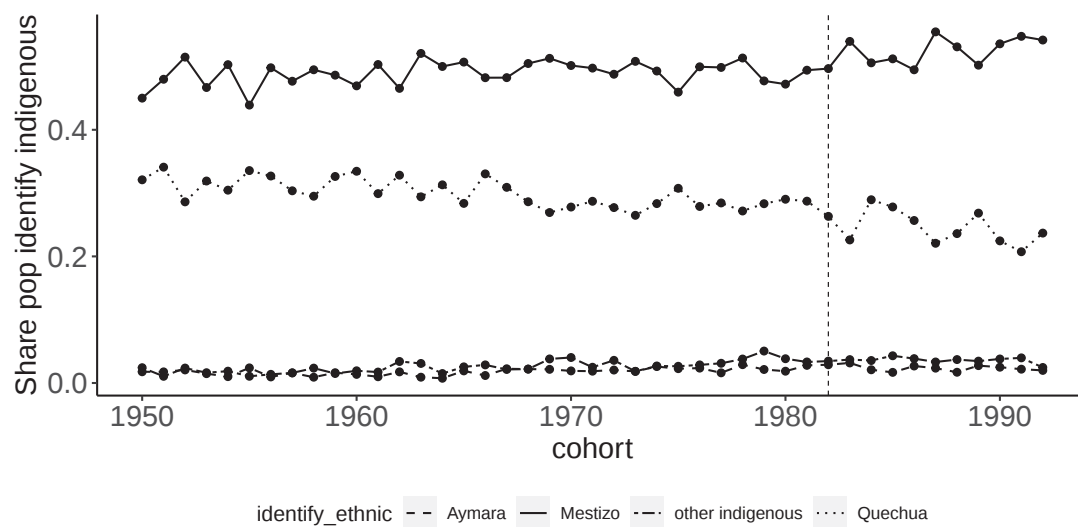
(b) Identify Indigenous

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. We 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.

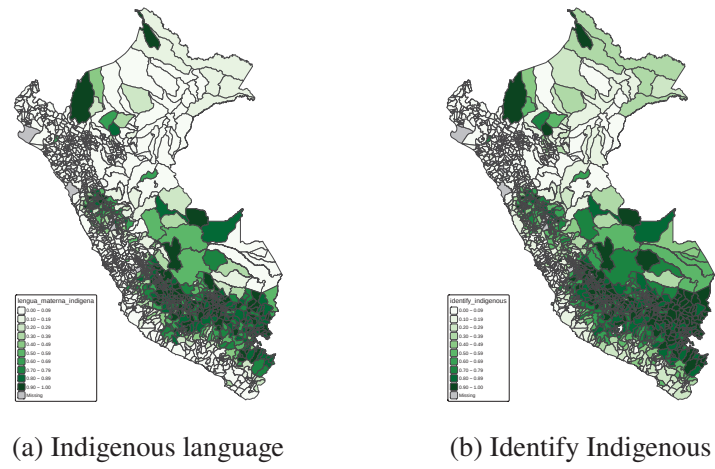


Figure A.3. Indigenous Identity

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.

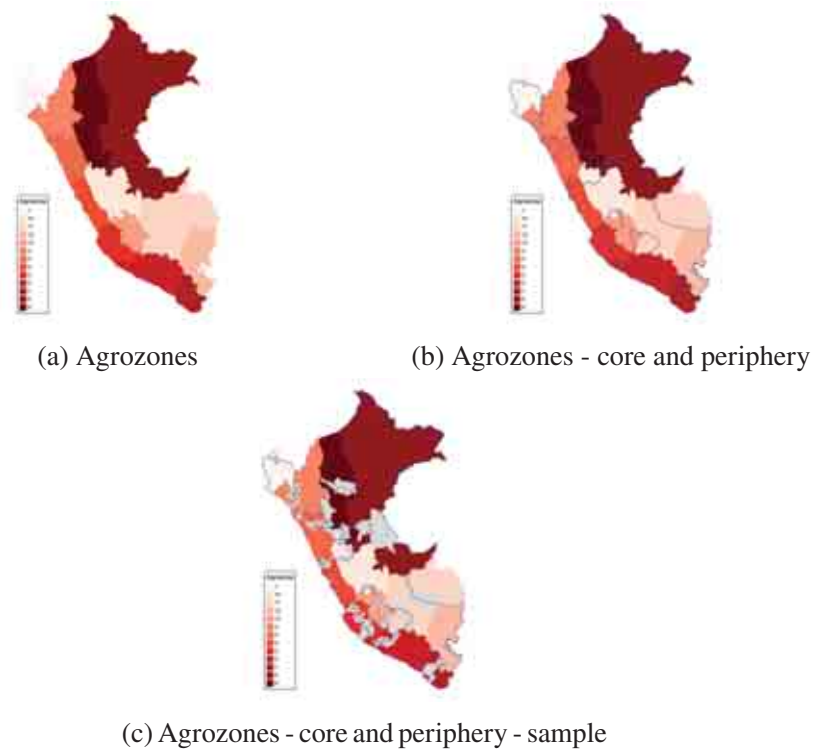


Figure A.4. Agrozones

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
<i>Full Sample</i>					
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
<i>Border Sample</i>					
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
<i>Border Sample variables</i>					
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** (0.007)	−0.034* (0.018)	−0.005 (0.005)	−0.023** (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*** (0.010)	−0.046 (0.030)	−0.009 (0.006)	−0.022* (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** (0.016)	-0.034** (0.016)	-0.050 (0.037)	-0.023** (0.009)	-0.023** (0.009)	-0.035** (0.014)		0.006 (0.004)
Migration		0.007 (0.004)	0.001 (0.008)		0.008 (0.008)	0.003 (0.011)		
Migration x Exposure			0.004 (0.017)			0.008 (0.020)		
Exposure violence x T x Migration			0.025 (0.039)			0.018 (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** (0.015)	0.002 (0.002)	-0.007* (0.004)	-0.011 (0.007)	-0.007 (0.007)	-0.012** (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*** (0.017)	−0.034 (0.029)	−0.017 (0.031)	−0.001 (0.031)
Exposure cohort 0-4	−0.088*** (0.010)	−0.070*** (0.023)	−0.020 (0.018)	−0.040** (0.018)
Exposure cohort 5-9	−0.049*** (0.008)	−0.044** (0.019)	−0.007 (0.020)	−0.025 (0.020)
Exposure cohort 10-14	−0.045*** (0.006)	−0.029 (0.022)	−0.014 (0.020)	−0.017 (0.020)
Exposure cohort 15-19	−0.014*** (0.005)	−0.006 (0.016)	0.009 (0.013)	0.008 (0.013)
Exposure cohort 25-29	0.007 (0.011)	0.036* (0.021)	0.004 (0.024)	0.023 (0.024)
Exposure cohort 30-34	0.003 (0.017)	−0.015 (0.060)	0.012 (0.035)	0.020 (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
<i>N</i>	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*** (10.868)	53.254** (22.093)	0.021 (0.016)	-0.067 (0.053)	-0.026 (0.029)	-0.021 (0.019)	-0.977** (0.471)
<i>N</i>	229	229	229	229	229	229	229
<i>R</i> 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 (0.101)	0.533 (0.740)	22.560 (23.592)	-0.428** (0.168)	0.119 (0.193)	3.861 (2.932)	1.355 (1.365)
<i>N</i>	229	229	229	229	229	229	229
<i>R</i> 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* (0.101)	-0.147 (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 (0.018)	0.013 (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 (0.020)	-0.003 (0.015)	0.035 (0.022)	-0.013 (0.020)	-0.005 (0.007)	0.012 (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* (0.006)	-0.012 (0.011)	0.011* (0.006)	-0.002 (0.014)	-0.007* (0.004)	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. Variable Definitions and Data Sources

Data sources

ENAHO. *Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza*, the Peruvian National Household Survey, administered annually by the *Instituto Nacional de Estadística e Informática* (INEI). We use the 2013–2017 waves.¹¹ Referred to below as ENAHO.

TRC. The database of the Peruvian Truth and Reconciliation Commission (*Comisión de la Verdad y Reconciliación*), which records 36,019 unique violent events reported in 16,917 testimonies collected between 1980 and 2000.¹² Referred to below as TRC.

LAPOP. The AmericasBarometer survey administered by the Latin American Public Opinion Project at Vanderbilt University. We use the 2006, 2012, 2014, and 2016 Peru waves.¹³ Referred to below as LAPOP.

Latinobarómetro. The annual Latinobarómetro Corporation public opinion survey. We use the 2002–2015 Peru waves.¹⁴ Referred to below as Latinobarómetro.

Albertus (2020) district data. Pre-conflict and conflict-period district characteristics used in the border-sample design are drawn from the replication data of Albertus (2020), which in turn compiles information from the 1961 and 1972 Peruvian censuses, the Peruvian land reform archives, and the TRC.

Unless otherwise indicated, individual-level variables restrict the sample to respondents born between 1958 and 1992. District-level variables from the border-sample design (Section D.1) use the 229 districts in our matched border sample.

C.1. Individual-level variables

Main outcomes

Identify as Indigenous.

Binary indicator equal to one if the respondent answers “Quechua,” “Aymara,” or “other Indigenous community” to the ENAHO question “According to your ancestors and your customs, do you consider yourself: Quechua, Aymara, other Indigenous community, Black, White, Mestizo, or unknown?”; and equal to zero if the respondent answers “Mestizo.” Respondents answering “Black,” “White,” or “unknown” are dropped from this variable. Source: ENAHO.

¹¹ <https://www.gob.pe/institucion/inei/informes-publicaciones/3364523-encuesta-nacional-de-hogares-enaho-2022>

¹² <https://sites.google.com/a/pucp.pe/informe-final-de-la-cvr--peru/>

¹³ <https://www.vanderbilt.edu/lapop/peru.php>

¹⁴ <https://www.latinobarometro.org/>

Indigenous mother tongue.

Binary indicator equal to one if the respondent answers “Quechua,” “Aymara,” or “another Indigenous language” to the ENAHO question “What is the language or mother tongue you learned in your childhood?”; and equal to zero if the respondent answers “Castellano” (Spanish). Respondents answering any other language (English, Portuguese, another foreign language, sign language, or other) are dropped from this variable. Source: ENAHO.

Treatment and exposure variables**Exposure.**

Binary indicator equal to one if the respondent’s district of birth experienced at least one violent event in which the victim was Indigenous during 1980–1992, and zero otherwise. Source: TRC, merged to ENAHO by district of birth and year of birth.

Exposure to non-Indigenous violence.

Binary indicator equal to one if the respondent’s district of birth experienced at least one violent event in which the victim was not Indigenous during 1980–1992, and zero otherwise. Source: TRC, merged to ENAHO by district of birth and year of birth.

Formative.

Binary indicator equal to one if the respondent was between 0 and 19 years old when the district of birth first experienced a violent event during 1980–1992, and zero if the respondent was 20 years or older at that time. Constructed from district and year of birth (ENAHO) and event year (TRC).

Formative (non-Indigenous violence).

Binary indicator equal to one if the respondent was between 0 and 19 years old when the district of birth first experienced a non-Indigenous-victim violent event during 1980–1992, and zero if the respondent was 20 years or older at that time. Used in the non-Indigenous-violence placebo test. Source: TRC and ENAHO.

Cohort indicators D_{ic} .

Categorical indicators for the respondent’s five-year age bin at the time of the first violent event in the district of birth, with categories $c \in \{\text{post}, 0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34\}$.

The “post” category flags respondents born in their district after the last recorded violent event there. The 20–24 bin is the omitted baseline in Equation 2.

Mechanism variables

Indigenous diversity index.

The probability that two randomly drawn Indigenous individuals from district d belong to different Indigenous groups, constructed following Alesina et al. (1999):

$$\text{Indigenous diversity index}_d = 1 - \sum_{i \in \mathcal{G}} (\text{Indigenous}_{id})^2, \quad (3)$$

where $\mathcal{G} = \{\text{Quechua, Aymara, other Indigenous}\}$ and Indigenous_{id} is the share of district d ’s Indigenous population that self-identifies as group i . To avoid contamination by the conflict itself, the index is constructed using only ENAHO respondents born before 1950—cohorts whose ethnic self-identification was established before the Shining Path’s first armed action in 1982. Source: ENAHO.

Indigenous diversity dummy.

Binary indicator equal to one if the Indigenous diversity index is strictly greater than zero (more than one Indigenous group is present in the district), and zero otherwise. Source: ENAHO.

Interpersonal trust.

Reverse-coded response to the LAPOP item IT1: “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?” Takes values 1 (untrustworthy), 2 (not very trustworthy), 3 (somewhat trustworthy), 4 (very trustworthy). Higher values indicate higher trust. Source: LAPOP.

Generalized trust.

Binary indicator equal to one if the respondent answers “You can trust most people” to the Latinobarómetro question “Generally speaking, would you say that you can trust most people, or that you can never be too careful when dealing with others?”; and zero if the respondent answers “One can never be too careful when dealing with others.” Source: Latinobarómetro.

Institutional trust.

Response on a four-point scale (1 = no trust, 2 = little trust, 3 = enough trust, 4 = a lot of trust) to ENAHO module 85's question "Currently, do you have trust in the following institutions?" We construct separate variables for trust in (i) the district government, (ii) the national government, (iii) the judiciary, (iv) the national congress, and (v) the Catholic Church. Source: ENAHO.

National identity (ordered).

Response to the Latinobarómetro question "How proud are you to be Peruvian?" on a four-point scale: 1 (not proud), 2 (little proud), 3 (proud), 4 (very proud). Source: Latinobarómetro.

National identity (dummy).

Binary indicator equal to one if the respondent answers "proud" or "very proud" to the Latinobarómetro national-pride question, and zero if the respondent answers "little proud" or "not proud." Source: Latinobarómetro.

Class-identity indicators.

Four binary indicators, each equal to one if the respondent reports current membership in, respectively, (i) a workers' association, (ii) a professionals' association, (iii) a political party organization, or (iv) a peasant community, and zero otherwise. Source: ENAHO.

Migration.

Binary indicator equal to one if the respondent's current district of residence at the time of the ENAHO interview differs from the district of birth, and zero otherwise. Source: ENAHO.

Individual controls**Age.**

Respondent's age in years at the time of the ENAHO interview. Source: ENAHO.

Year of birth.

Respondent's calendar year of birth, bounded to 1958–1992 by the sample restriction. Source: ENAHO.

Sex.

Binary indicator equal to one if the respondent is female, and zero if male. Source: ENAHO.

D. Construction of the Co-Local Perpetrator Statistic

Data source

We compute the co-local perpetrator statistic reported in Section 4.1 from the event-level file (`actos_est.sav`) of the Peruvian Truth and Reconciliation Commission's *Base de Datos Estadística*. Each row of this file corresponds to one victim of one violent act. Multiple rows sharing the same event identifier (`IDEVENTO`) represent different victims or actors in the same event.

Variables

The following TRC variables are used:

[`noitemsep`, `topsep=2pt`]

- `E02`: binary indicator for acts perpetrated by the PCP (Sendero Luminoso); `E02 = 1` identifies Shining Path perpetration.
- `IDIOMA`: victim's mother tongue, coded 1 = Spanish, 2 = Quechua, 3 = other native Peruvian language.
- `UBIGEONA`: 10-digit UBIGEO code identifying the victim's district, province, and department of birth.
- `UBIDIST`, `UBIPROV`, `UBIDEP`: UBIGEO codes identifying the district, province, and department where the violent event occurred.

Sample construction

We restrict the TRC event-level data to Shining Path perpetration against Indigenous victims. From 36,019 total victim-act records, 14,186 correspond to Shining Path perpetration (`E02 = 1`). Of these, 7,674 record an Indigenous mother tongue (`IDIOMA` $\in \{2, 3\}$) for the victim, covering 3,609 unique events. Mother tongue is recorded for 100 percent of Shining Path victim-acts, of which 54.1 percent are Indigenous-language speakers. This share describes the actor-level victim-act file (`actos_est.sav`) and is not directly comparable to the seventy-five percent share of fatal victims of the conflict reported in the TRC's *Informe Final* (Laplante and Theidon, 2007): the latter refers to deaths and disappearances across all perpetrators, while the former refers to victim-acts of Shining Path violence across all categories of violent acts (killings, disappearances, illegal detentions, kidnappings, forced recruitments, torture, rapes, and woundings). Victim birthplace at district

precision is available for 94.1 percent of the Sendero Indigenous sample; at province precision, for 95.3 percent.

Construction of the matching statistic

The TRC records the geographic location of each event and the birth location of each victim but does not record the geographic origin of the perpetrators. The statistic we report is therefore a *locality-of-violence* measure: the share of Shining Path events against Indigenous victims in which the violent act occurred in the victim's own birth district (or province, or department). Under the assumption, documented in qualitative TRC testimony and in the literature on the Shining Path's operational structure, that its cells recruited locally and operated within the communities where members lived, a high locality-of-violence share implies a high share of co-local and, by extension, co-ethnic perpetration.

Formally, we extract the first six, four, and two digits of UBIGEONA to obtain the victim's birth district, province, and department codes respectively. Records for which UBIGEONA indicates only province or department precision (identified by trailing zeros in positions 5–6) are excluded from the district-level denominator but retained at coarser administrative levels where appropriate. Matching is defined as exact equality between the extracted birthplace code and the corresponding event-location code (UBIDIST, UBIPROV, UBIDEP).

Results

Table A.12 shows the main statistics. Of the 7,223 Shining Path victim-acts in the actor-level file with an Indigenous mother tongue and birthplace recorded at district precision, 63.6 percent took place in the victim's own birth district. The corresponding shares at province and department precision are 74.8 percent (5,470 / 7,314) and 83.6 percent (6,414 / 7,674) respectively. Collapsing to unique events yields nearly identical figures: 63.1 percent at the district level (2,195 / 3,481), 73.0 percent at the province level (2,577 / 3,531), and 85.6 percent at the department level (3,088 / 3,609). The reported figures are conservative lower bounds: migration between birth and event location by either the victim or the perpetrator before the conflict would attenuate the observed match rate, even if all perpetrators were truly co-local.

Table A.12. Shining Path violence against Indigenous victims: geographic match between event location and victim's birth district

Administrative level	Victim-act level		Event level	
	Matches	Share	Matches	Share
District	4,596 / 7,223	63.6%	2,195 / 3,481	63.1%
Province	5,470 / 7,314	74.8%	2,577 / 3,531	73.0%
Department	6,414 / 7,674	83.6%	3,088 / 3,609	85.6%

Notes: The table reports the share of Shining Path violent events against Indigenous victims where the event location matches the victim's birth location, at three levels of geographic precision. The victim-act level treats each victim within each event as a separate observation; the event level collapses multiple victims within the same event into a single observation (identified by IDEVENTO). Denominators at the district and province levels exclude records for which the victim's birth location is recorded at coarser precision (see Appendix D for construction details). Data from the TRC *Base de Datos Estadística*, actor-level file `actos_est.sav`.

D.1. Border-sample district variables

These variables are used in Figure 2 to test pre-conflict balance between core and periphery districts in the border sample. Unit of observation is the district. All variables are drawn from the replication data of Albertus (2020).

Conflict outcomes (1980–2000)

Violent events.

Number of violent events recorded in the district during the Peruvian civil conflict, 1980–2000, including murders, disappearances, illegal detentions, kidnappings, forced recruitments, torture, rapes, and woundings.

Deaths.

Number of deaths recorded in the district during the Peruvian civil conflict, 1980–2000.

Pre-conflict political characteristics

Marxist vote share.

Share of votes cast for Marxist political parties in the 1980 presidential election.

Social movements.

Number of autonomous communal uprisings recorded in the district prior to the Shining Path conflict.

Pre-conflict demographic and economic characteristics**Non-Indigenous population, 1961.**

Share of district population reporting Spanish (Castellano) as their mother tongue in the 1961 census.

Illiteracy rate, 1961.

Share of district population unable to read or write in the 1961 census.

Illiteracy rate, 1972.

Share of district population unable to read or write in the 1972 census.

Population, 1972.

District population in thousands in the 1972 census.

State employees, 1961.

Number of state personnel recorded in the district in the 1961 census.

Land and geography**Land reform hectares (per capita).**

Hectares per capita redistributed through the 1969–1975 Peruvian land reform in the district.

Private land (hectares).

Thousands of hectares of privately held land in the district.

Share of cultivated land.

Share of district area under cultivation.

Elevation.

District mean elevation, in kilometers.

Slope.

District mean slope, in degrees.

E. Comparison with Albertus (2020)

The geographic basis for our border-sample design, Peru’s agrarian reform zones and their core–periphery boundaries, follows Albertus (2020), who shows that core districts received substantially more land redistribution than periphery districts and that this difference in land reform is associated with differences in subsequent Shining Path conflict intensity. Our paper uses the same underlying geographic variation but addresses a different question with a different identification strategy. Three distinctions are worth drawing out explicitly.

Question. Albertus studies the effect of land reform on conflict. We study the effect of conflict on ethnic identity. The agrarian reform zones are, for Albertus, a source of variation in the treatment of interest; for us, they are a sample-construction device that ensures we compare districts similar on pre-conflict observables but differentially exposed to violence.

Design. Albertus estimates a regression discontinuity design, using the core–periphery boundary as a running variable with treatment status “jumping” at the threshold. We use the boundary as a sample restriction: we retain only districts within a narrow band of the boundary and estimate a difference-in-differences specification exploiting cross-cohort variation in age at exposure to violence. Our approach is closer in structure to Lichter et al. (2021), who use German district boundaries to identify long-run effects of state surveillance. The key consequence of this design difference is that our source of identification is not the level of violence across districts but the *timing* of violence within districts relative to individual cohorts’ identity formation windows.

Zones included. Albertus restricts his sample to agrarian reform zones 1 through 7 and 11, where land reform was consistently implemented and the core–periphery structure was preserved through the entire reform period. We include all thirteen zones. Albertus’s exclusions are: Zones 8 and 9 (minimal land reform applied); Zone 10 (split into Zones 10 and 13 in 1974, complicating the core–periphery classification); and Zone 12 (never classified into core and periphery). Because our identification strategy does not require a clean running variable at the boundary, but only a set of reasonably comparable districts on either side, these complications are less consequential for us than for the RDD.

The inclusion of additional zones may explain the difference between Albertus’s main finding

that core districts experience *less* conflict and our Figure 2 result that core districts in our sample experience *more* conflict. Two possibilities, neither of which we can test directly in this paper, are consistent with this pattern. First, in Zones 8 and 9, land reform was low-intensity, which some of the land-reform-and-conflict literature associates with *more* conflict rather than less—either because recipients mobilize for further redistribution (Albertus and Kaplan, 2013), or because weak reform leaves pre-existing grievances intact without providing the compensating state presence that would deter insurgent recruitment. Second, Zone 10 in Junín and Zone 13 in Pasco lie on the northern edge of the central Andean highlands, a region in which Shining Path operations extended substantially after the initial Ayacucho-centered phase of the insurgency. The local state presence in these departments was limited during the conflict’s early years (CVR, 2003), which may have contributed to the observed pattern of violence being more concentrated in core districts in our expanded sample. These are hypotheses rather than findings.

We emphasize that neither the direction nor the magnitude of our main identification result depends on the interpretation of this discrepancy. What matters for our analysis is that, within the border sample, core and periphery districts are balanced on pre-conflict characteristics that predict Indigenous identity (Figure 2), and that violence intensity varies across them. The direction of that variation, whether core or periphery receives more, is not itself an input to our difference-in-differences estimates, which exploit within-district variation across cohorts.

Related literature on land reform and conflict. For completeness, we note that the broader literature on land reform and conflict has reached mixed conclusions, with some studies finding a negative relationship (Albertus, 2020; Guardado, 2018; Huntington, 2006; Kalyvas, 2006; Wood, 2003) and others a positive relationship (Albertus et al., 2018; Albertus and Kaplan, 2013; Finkel et al., 2015; Mason, 1998). Proposed mechanisms for the negative relationship center on opportunity cost: land-receiving peasants face a higher cost of joining an insurgency (Guardado, 2018). Proposed mechanisms for the positive relationship include low-intensity reforms that mobilize rather than pacify recipients (Albertus and Kaplan, 2013) and withdrawal of landowner-provided local infrastructure (irrigation, market access) without adequate state replacement (Kapstein, 2017). Because land reform is not our treatment of interest, we do not take a position on which mechanism dominates; we note the literature only to contextualize the departure from Albertus’s conflict-sign result in our expanded sample.