The political economy of socioenvironmental conflict: Evidence from Peru*

David Kreitmeir†

October 19, 2024

Abstract

Over the past two decades, violence against land and environmental activists has been on the rise, besetting even stable democracies. Using a unique, fine-grained data set on social conflict events in Peru and exogenous variation in world mineral prices, I document a strong link between local mineral rents and violent state repression of socioenvironmental protests in a democratic institutional setting. I show that the increase in the use of excessive force cannot be explained by changes in protester behavior. Empirical findings highlight the role of local authorities: the election of a pro-mining mayor is associated with a higher prevalence of state repression and corruption in the constituency. The legal and democratic accountability of local authorities is, however, found to be limited. The reported increase in corruption does not translate into more investigations against pro-mining mayors for corruption offenses nor are reelection results of incumbents found to be negatively affected by state violence against protesters. Finally, I show that violent state repression is successful in forestalling conflict resolution agreements that acknowledge protesters' demands.

Keywords: Resource curse, mining, social conflicts, local authorities, Peru
JEL Codes: D74, H7, O13, O16, P16, Q34.

*I thank Sascha O. Becker, Paul A. Raschky, Thomas Ueberfuhr, Alexander Ballantyne, Ashani Amarasinghe, Lukas Wellner, the SoDa Labs PhD group, and participants of the Melbourne PolEcon/PolSci meeting, the 13th Australasian Public Choice Conference, Monash Environmental Economics Workshop, Australasian Development Economics Workshop, 10th Annual Workshop on Natural Experiments, and the Australian Conference of Economists for their helpful suggestions and comments.

†Monash University, Department of Economics, 900 Dandenong Rd, Caulfield East, VIC 3145, Australia. Email: david.kreitmeir1@monash.edu.

1 Introduction

From March 2004 to December 2019, more than 386 anti-mining activists were injured and at least 41 killed in Peru during demonstrations against expropriation of local lands and inadequate compensation for environmental damages.[1] This violence against land and environmental defenders is not isolated to Peru: it is part of a worrying global trend of endemic violence against socioenvironmental activists besetting even stable democracies such as Peru.[2] Juxtaposing the number of land and environmental defenders killed during the 2002–2019 period with the number of civil conflict fatalities in Figure 1, reveals the relative prevalence and severity of socioenvironmental conflicts, particularly in mineral-rich and democratic countries in Latin America and Southeast Asia. This form of conflict has so far, however, received relatively little attention in the resource curse literature, which has primarily focused on armed civil conflict over the appropriation of natural resource rents, predominantly in weakly institutionalized countries (e.g. Dube and Vargas, 2013; Berman et al., 2017; Sánchez De La Sierra, 2020).

Utilizing a unique granular data set on violence against protesters during socioenvironmental conflicts in Peru and exogenous variation in monthly world mineral prices, this paper provides new empirical evidence on the causal relationship between mineral rents and violent state repression in a democratic institutional setting: a rise in local mineral rents is estimated to increase the propensity of both nonfatal and fatal violence against protesters opposing mining projects.

Peru's institutional framework provides a unique setting to shed light on the nexus between mineral rents and the suppression of socioenvironmental protests.[3] Municipality governments in Peru are highly dependent on central government transfers due to their limited ability to levy taxes, but handle more than 20% of the total national government budget and are responsible for more than 40% of public investments (Pique, 2019). Two of

[1] These figures are the author's calculations based on reports published by the Peruvian Office of the Ombudsman.

[2] Peru exhibited a *polity2* score (Marshall et al., 2002) of 9 (of a maximum 10 and a minimum -10) throughout the period from 2002 to 2018. In relative terms, Peru's *Polity2* score equaled that of France over this period and surpassed that of the United States from 2016 to 2019.

[3] Figure A.2 in the Appendix graphs the evolution of mineral rents and number of social conflicts related to formal mining in Peru for the 2004-2019 period.

the most lucrative transfers available to municipalities are based on local mining activity. According to a fixed allocation rule, the central government transfers 50% of income taxes collected from mining companies and 80% of royalties to municipal governments in mining regions. While recent studies have shown that this redistribution scheme can positively affect human capital accumulation (Agüero et al., 2021) and living standards in mining municipalities (Loayza and Rigolini, 2016), it can also create adverse incentives for local authorities to repress opposition to advance local mining projects.

In this paper, I combine information on mineral production and concessions at the municipality level with social conflict events to test whether an increase in prospective mineral rents raises the propensity of violence against activists during socioenvironmental protests. My identification strategy relies on exogenous changes in world mineral prices inducing variation in prospective mineral rents accruing to local governments . I find that a rise in mineral prices increases the probability of observing protester arrests, injuries and deaths during confrontations with police forces. Leveraging the spatial and temporal granularity of the data set, I conduct the analysis at the month and municipality level, the third and lowest administrative level in Peru. By conditioning on municipality $\times$ year fixed effects, I account for time-varying and time-invariant municipality characteristics that could confound the estimated relationship such as changes to municipal government funds or economic development. I show that the estimated positive relationship between mineral rents and violence against protesters cannot be explained by behavioral changes of activists. Specifically, I show the absence of a systematic relationship between mineral prices and protester riots or violence against police officers. Moreover, benchmark estimates are robust to the exclusion of months with protester riots and across numerous sensitivity tests.

Having documented how mineral rents affect state repression of socioenvironmental protests, I turn to the role of mayors who preside over municipal governments and are in charge of local security and the provision of other basic public goods. Using a continuity-based regression discontinuity design (RDD), I test whether the election of a *pro-mining* as opposed to an *anti-mining* candidate has a positive effect on the prevalence of both, police

violence against activists and corruption in the municipality. To do so, I collect a new data set on mayoral government plans published ahead of elections. Using natural language processing (NLP) and large language models (LLMs), I label over $9,000$ candidates on basis of their revealed sentiment toward formal mining activities in government plans. I find empirical evidence in support of the hypothesis that the narrow election of a *pro-mining* candidate positively affects police violence during their term. Further, the RDD results are indicative of political capture of the judicial process. The election of a *pro-mining* candidate is estimated to increase the number of corruption incidences reported by constituents. This increase is, however, not reflected in a coincident rise of investigations against *pro-mining* mayors for corruption offenses in office. Rather, estimates suggest that *pro-mining* mayors are more effective in stalling initial investigations against them.

Next, I examine how the incumbency advantage is affected by state repression during the mayor's term to gauge democratic accountability in local elections. Using a modified RDD that compares the incumbency advantage of previous election winners vis-à-vis runner-ups (de Benedictis-Kessner, 2018; Lewis et al., 2020), I find that reelection results of incumbents whose terms were characterized by violent state repression are indistinguishable to those of mayors whose terms were not marred by violence against protesters. Overall, the absent negative impact of state violence on the incumbency's relative reelection performance indicates limited political accountability of mayors.

In the last part of the paper, I provide some first empirical evidence on how violent state repression affects the final social conflict outcome. Since standard regression models fail to appropriately adjust for time-dependent observed confounders that are themselves affected by previous incidences of police violence, I rely on incremental interventions that shift the propensity score (Kennedy, 2019) to estimate the causal effect of state repression on the final outcome. In contrast to other dynamic causal inference methods such as marginal structural models (e.g. Blackwell, 2013; Imai and Ratkovic, 2015), this estimator is completely nonparametric and avoids strong statistical assumptions about the treatment process. Exploiting the unique information on the evolution of social conflicts over time in the data set, I show that, while violence against protesters prolongs

the duration of the social conflict, state repression is successful—from the perspective of local authorities—in bringing social conflicts to a standstill without making costly concession to the opposition.

This study relates to several strands of the literature. First, by moving the focus from armed civil and political conflict over the control of natural resource rents (s., e.g., Blair et al., 2021, for a review) to *socioenvironmental* conflict, the study contributes to the rich empirical literature on the nexus between natural resource abundance and conflict. Dube and Vargas (2013)'s seminal study compares the effect of positive oil price shocks and negative coffee price shocks on paramilitary attacks in Colombia, finding evidence for both a rapacity and an opportunity cost effect. Using granular data on local mining activity, Berman et al. (2017) study mineral resource fueled conflict in Africa. The central role of mineral rents in financing armed conflict is also emphasized in recent work by Sánchez De La Sierra (2020), who finds that rebel groups in the DRC establish quasi-states around mines, and by Limodio (2022), who exploits exogenous movements in silver prices to examine the extent of terrorism finance in Pakistan. By contrast, positive price shocks to agricultural commodities have been shown to increase the opportunity cost to fighting and are negatively related to armed conflict (Bazzi and Blattman, 2014; Berman and Couttenier, 2015; McGuirk and Burke, 2020; Blair et al., 2021). This paper documents a *new type of resource conflict* concerned with the protection of communal lands from negative externalities of mining. It highlights the violence potential of mineral rents independent of armed rebellion or agricultural income shocks. Granular information on social conflicts and confrontations between police forces and protesters allows me, additionally, to discern whether the observed state violence is a response to changes in protester behavior or the result of strategic use of excessive force.

Second, the study's results contribute to a growing strand of the literature exploring the determinants of land and environmental conflict (Butt et al., 2019; Haslam and Ary Tanimoune, 2016; Grasse, 2022) and how institutions influence social and resource-related conflicts (Arellano-Yanguas, 2011; Fetzer and Marden, 2017; Sexton, 2020). In particular, it augments theoretical and empirical findings of Besley and Persson (2011)

and Fetzer and Kyburz (2024) that cohesive political institutions discourage political violence by providing causal evidence of the relationship between mineral rents and state repression in the context of a strong democracy. My findings highlight the importance of considering decentralization and fragmented, candidate-centered local level politics when designing institutions to break the resource course.[4] More generally, the study informs previous theoretical and empirical work on *extractive institutions* (Acemoglu et al., 2001; Acemoglu and Robinson, 2012). Exploiting time series variation in the value of resources in a controlled institutional setting, I provide new empirical insights into how extractive institutions may persist at the local level notwithstanding democracy. Furthermore, the study speaks to theoretical work on the importance of considering the influence of powerful corporations on political institutions (Zingales, 2017) and the role of civil society in constraining an otherwise overreaching state (Acemoglu and Robinson, 2020).

Third, this paper relates to the literature on the impact of natural resource windfalls on local communities and politics. Revenue sharing of resource rents across different tiers of government has been found to improve education and the standard of living in local communities (Agüero et al., 2021; Aragón and Rud, 2013; Loayza and Rigolini, 2016; Maldonado and Ardanaz, 2022) but has also been connected to inefficient use of public funds (Maldonado and Ardanaz, 2022), corruption (Maldonado, 2011; Baragwanath Vogel, 2021), and negative selection of candidates running for and elected into office (Brollo et al., 2013; Carreri and Dube, 2017). Asher and Novosad (2023), moreover, show that local mining booms increase criminal politicians' likelihood of being elected and committing crimes in office. I complement these findings by providing a comprehensive analysis of the nexus between mineral rents, state repression, and local political institutions.

Lastly, the paper makes a methodological contribution. It provides some first empirical evidence of how state repression impacts the final conflict outcome using the empirical framework of *incremental interventions on the propensity score* developed by Kennedy (2019). Incrementally shifting the probability of treatment (i.e. state violence against

[4]Peru exhibited a *xconst* of 7 (of a maximum 7 and a minimum 1) and a *polity2* score of 9 (of a maximum 10 and a minimum -10) throughout the period from 2002 to 2018 (Marshall et al., 2002), the preferred indicators of Besley and Persson (2011) to measure cohesive political institutions.

protesters) allows me to adjust for time-varying confounders affected by and affecting the treatment decision without having to impose the strong statistical assumptions on the treatment process required in *marginal structural models (MSM)* that have primarily been used in the literature to identify dynamic causal treatment effects in longitudinal studies (Blackwell, 2013; Imai and Ratkovic, 2015; Ladam et al., 2018; Bodory et al., 2022, among others).

2 Mineral Rents and Social Conflict Violence

2.1 Institutional Framework

Peru has three administrative levels, with groups of municipalities (*distritos*) forming provinces (*provincias*) and groups of provinces forming regions (*departamentos*). In 2001, as part of a fiscal decentralization reform, the national government implemented a fiscal transfer scheme—the so-called *canon minero*—which distributes 50% of the income taxes collected from mining corporations back to local governments according to a fixed allocation rule. In particular, since the enactment of of Law No. 28327 in 2004, 10% of this amount is directly distributed to producing municipalities; 20% of the amount is shared among all municipalities in a producing province, and 40% is distributed among all municipalities in a producing region. In addition, Law No. 28258 in 2004 established that 20% of the collected mining royalties are directly transferred to the municipality(s) where the mining concession is located, 20% are distributed among municipalities in the province of the mining concession, and 40% are allocated to municipalities in the region of the concession.[5] Table B.1 in the appendix presents summary statistics for the amount of total revenues in real 2010 USD received by municipalities from either the *canon minero* or royalties over the period 2004 to 2020. Additionally, the total revenues are disaggregated by municipality category, i.e., whether a municipality had mining production, only concessions but no active production, or neither during the sample period. Table B.1

[5]With the exception of the 10% of the *canon minero* and the 20% of the mining royalties directly transferred to municipalities, the exact allocation of *canon* transfers and royalties to municipalities is determined by social and economic characteristics in the respective municipality.

illustrates the favorable position of mineral-producing municipalities and the substantial financial incentives for local governments to attract mining projects. In particular, local governments are highly dependent on fiscal transfers from the central government, with the share of the *canon minero* accounting for approximately 29% of local governments' budgets and as much as 70% of those of producing municipalities (Canavire-Bacarreza et al., 2012; Maldonado and Ardanaz, 2022).[6] Yet, the mayor as the municipality's administrative authority handles the local budget of more than 20% of the total government budget. Moreover, municipalities are responsible for more than 40% of the country's total public investment (Pique, 2019). While the use of funds received from the *canon minero* and royalties is tied to specified objectives, the limited impact of transfers on living standards has been connected to corruption and unproductive but politically favorable investments in local infrastructure and public employment (Maldonado, 2011; Crabtree, 2014; Maldonado, 2017). In particular, the electoral system, which grants the elected mayor 50%-plus-one seats in the municipality council independent of the election results, has been identified to severely limit political accountability in the absence of a strong civil society (Maldonado, 2011; Crabtree, 2014).

2.2 Data

2.2.1 Social Conflicts

For this study, I construct a unique data set of social conflicts related to mining projects in Peru based on reports by the Office of the Ombudsman (*Defensoría del Pueblo*), henceforth referred to as "the Ombudsman."[7] Since April 2004, the Ombudsman has published monthly reports on social conflicts in Peru. These reports follow a fairly uniform format, which enables me to identify and track developments of individual social conflicts over time. In particular, each report entry constitutes a separate social conflict and

[6]The dependence of local governments on central transfers in Peru is the result of their marginal ability to levy and collect taxes, with self-raised revenues contributing on average only 12.5% to the local budget (Aresti, 2016).

[7]Other empirical studies have used Ombudsman reports for information on social conflict incidence (Arellano-Yanguas, 2011; Haslam and Ary Tanimoune, 2016; Castellares et al., 2017; Orihuela et al., 2019; Sexton, 2020, among others). However, no study has so far used the social conflict data at such fine-grained spatial and temporal resolution.

contains the following key information: type of social conflict, location, actors, current status (active, latent, resolved or removed) and a description of recent events related to the conflict if any transpired.[8] An example entry is presented in Figure A.1 in the Appendix.

I restrict the set of social conflicts for this study to conflicts concerning industrial mining.[9] Therefore, I exclude conflicts related to informal or illegal mining activities to account for fundamental differences in the characteristics of these conflicts. In particular, the latter often take place in regions with no official mining production and—by definition—involve actors more prone to using illegal means to secure mining rents. While a role of local authorities as stakeholders in these illegal operations cannot be ruled out, the concern of this study is the "visible" suppression of mining opposition by police forces.[10] Note that this set of conflicts also comprises unfulfilled promises about employment for local community members in exchange for licensing their land but excludes labor disputes about inadequate pay, dangerous working conditions.

The baseline sample is, moreover, confined to social conflicts with location information at the municipality level to identify whether and how changes in local mineral rents lead to escalations in local anti-mining conflicts. The resulting data set is a panel at the conflict–month level tracking each social conflict from its start until its resolution.

I complement the information on the evolution of social conflicts with data on protest activities against mining activities on basis of the description of recent events in the Ombudsman reports. In particular, I code the date or month of the protest, the location of the event, and the number of arrests, injuries, and casualties among protesters and police forces.[11] I aggregate all variables to the monthly level to match the frequency of social conflict reports and rely on binary coding of all protest measures in the empirical analysis.

[8] If a conflict is inactive (latent) for an extended period of time, the Ombudsman removes the case from the register.

[9] I.e. one of the primary conflict parties has to be an industrial mining entity.

[10] In isolated instances, those confrontations can also feature private security guards.

[11] For cases in which the exact number of injuries or arrests is not clearly specified, the number is coded as one to provide a conservative estimate. Note that this coding decision does not affect the estimates in the analysis since all dependent variables are at the extensive not intensive margin to minimize potential measurement errors.

For the baseline analysis, I use the granular information on locations of protests and social conflicts to construct a balanced panel of social conflicts and their associated protests at the municipality–month level for the period from March 2004 to December 2019.

2.2.2 Mineral Production and Prices

I obtain data on monthly mineral production disaggregated by type of mineral at the municipality level from the Ministry of Energy and Mining (MINEM) for the period from 2002 to 2019.[12] I combine the information on mineral production with monthly world prices on minerals provided by the *World Bank*. The *World Bank Commodity Price Data* cover seven minerals that represent more than 99% of the total production value over this period.[13] All mineral prices are uniformly expressed as real USD per kilogram.

Following Berman et al. (2017), I determine the main mineral in a municipality on the basis of its total production value over the period 2002–2019 evaluated at mineral prices at the start of the period 2002. As an alternative price measure, I construct a weighted price index of all minerals mined in a municipality, with weights equal to each mineral's share in total production value in the municipality over the 2002–2019 period. In addition, I consider the presence of mining concessions in nonproducing municipalities. Data on mining claims are drawn from the *SNL Mining & Metals* database. This is in line with anecdotal evidence from the Ombudsman that social conflicts often break out before the production stage, when a mining concession is granted. In conjunction with royalty revenues being based on concession location rather than production value, an analysis relying exclusively on mineral production might not capture the "total" effect of changes in mineral rents on the violent suppression of activists. Figure 2 illustrates this nexus. A substantial share of the social conflicts associated with industrial mining are located in

[12]Note that MINEM also provides production data for the year 2001. However, the data are available only at the regional level. For more details, see: minem.gob.pe/_estadistica.php?idSector=1&idEstadistica=12501.

[13]The minerals are iron, copper, lead, tin, zinc, gold, and silver. Price timelines for each of those minerals are presented in Figure A.6 in the Appendix. Price changes in the top 75th (90th) percentile are depicted in blue (red). Note that I collect prices at the yearly level for the remainder of minerals mined in Peru from the USGS. For more details on mineral prices, see Appendix Section A.3.

municipalities with no mineral production during the study period. However, in many of these municipalities, mining concessions—visualized by white rectangular areas framed in black—were granted during this time. Figures A.3 and A.4 in the Appendix provide analogous images for mining royalties and *canon minero* distributed to each municipality during the 2002–2019 period. Since information on the selling price or projected value of concessions is not available in the *SNL Mining & Metals* database, I determine the main mineral for a nonproducing municipality on the basis of a simple count of each concession's primary commodity. For the weighted price index, weights are calculated based on the primary commodity counts of each mineral.

Table 1 provides descriptive statistics for the main variables of interest disaggregated by presence of mining production or concessions. The total data set comprises 1873 municipalities over 18 years; 228 municipalities had a positive production value over the 2002–2019 period, with 278 additional municipalities granting at least one mining concession. The probability of observing the violent death of an activist or use of force against activists in a municipality in a given month is between 0.01% and 0.04%. The likelihood of observing any of these events more than quadruples when I consider only municipalities with mining production. An increase in the average probability of observing protester arrests or injuries is also observed for municipalities with mining concessions granted during the study period.

2.2.3 Other Data

I obtain data on *canon minero* and mining royalty transfers from the Ministry of Economy and Finance (MEF).[14] I convert all yearly transfer payments into real USD using the official yearly exchange rate and MUV Index provided by the *World Bank*. Information on administrative boundaries is retrieved from the National Institute of Statistics and Information Technology (INEI).[15]

[14]The MEF makes the data publicly available through its Transparency Portal (*Transparencia Económica*), which can be accessed here: apps5.mineco.gob.pe/transferencias/gl/default.aspx.

[15]The data can be accessed here: datosabiertos.gob.pe/dataset/resource/a43e17c8-fa37-463d-aa7e-2ce2a272491b.

2.3 Empirical Strategy

Leaning on the empirical framework in Berman et al. (2017), the baseline linear probability model (LPM) takes the form:

$$A_{it} = \delta \left(M_i \times \ln(P_{it}^W) \right) + \mathbf{X}_{it}^{'} \beta + \gamma_{iy} + \epsilon_{it}, \tag{1}$$

where A_{it} takes on the value one if a forceful action against protesters is taken by the police force (or vice versa) in month t in municipality i and 0 otherwise. M_i equals one if there has been either *(i)* mineral production or *(ii)* a mining concession being granted in municipality i during the period 2002–2019, and zero otherwise. My use of a constant mining indicator variable accounts for potential reverse causality. In particular, local opposition to a mining project could lead to a halt in production or the withdrawal of the mining concession. Holding the production indicator fixed ensures that the coefficient of interest δ is identified by exogenous movements in world mineral prices only.

The inclusion of municipality $\times$ year fixed effects, γ_{iy}, accounts for time-invariant municipality characteristics simultaneously affecting social conflict and local mining activity such as property rights enforcement or historical disenfranchisement of indigenous communities. My focusing exclusively on within-municipality–year variation, moreover, lets me control for time-varying differences across municipalities at the yearly level, in particular economic and budgetary changes in municipalities. P_{it}^W denotes the world price of the *main* mineral in municipality i. For municipalities with no mining production (or municipalities with no concessions), the mineral price is set to zero. For my baseline estimates, the main mineral in a municipality is the mineral with either the highest total production value (in 2002 USD prices) or the highest primary commodity count among granted concessions. I test the robustness of my results derived with the baseline price variable to my use of an alternative value-weighted price index. In particular, this price index is the value-weighted price of all minerals mined in municipality i over the 2002–2019 period, with weights equaling the production value in producing municipalities. For nonproducing concession municipalities, the weights are determined on basis of the num-

ber of primary commodities of granted concessions in the municipality. Additionally, I test the sensitivity of my baseline estimates to my restricting the sample to only producing municipalities. $\mathbf{X}_{it}$ is a set of potential time-varying codeterminants of social conflict that I consider in robustness checks.

Given the granularity of the data and the spatial clustering of events, the standard errors are allowed to be spatially and temporally correlated. Specifically, I correct the standard errors for heteroskedasticity and autocorrelation (HAC), retaining a radius of 500 km for the spatial kernel.[16] I assume a linear decay in distance for the spacial correlation for nonzero elements in the HAC matrix.[17] Serial correlation is allowed to be "unconstrained" and can span the entire sample period. I also provide robustness results for alternative spatial radii.

2.4 Baseline Results

Table 2 presents the results for three baseline social conflict incidents involving use of force against activists during public protests by police forces.[18] I estimate that a one-percentage-point increase in the main mineral price increases the probability of protesters' being arrested by 0.04 percentage points, of their being injured by 0.19 percentage points, and of their being killed by 0.07 percentage points. These effects are sizable. Relative to the mean, a one-percentage-point increase in mineral prices more than doubles the probability of my observing use of excessive force.

Having established a positive relationship between projected mineral rents and use of force against protesters, columns 4 and 5 test whether the observed escalation of social conflicts is indeed the result of a harsher crackdown by police forces or can be explained by more militant behavior on activists' side provoking a reaction from police. While the conflicts under investigation are concerned with environmental and social issues related to mining activity and not the appropriation of mineral resources as considered in the civil conflict literature (Berman et al., 2017), frustrations over insufficient compensation

[16] The radius of 500 km is taken as it is close to the median internal distance of 540 km.

[17] In detail, the nonzero elements of the HAC spatial pattern matrix follow a Bartlett kernel (Colella et al., 2023).

[18] In rare instances, private security guards are involved in the confrontations.

for environmental damage or properties might surge among protesters when the expected rents for mining corporations and local authorities increase. To address this potential channel, I use information on injuries and killings of police officers during demonstrations to create a binary indicator that equals one if violence against police forces is observed and zero if otherwise. Note that this variable could also capture acts of retaliation at the hands of protesters in response to initial violence by police. The estimated effect might hence overstate the impact of changes in mineral rents on protester violence.

Reassuringly, I find no evidence for an effect of mineral prices on protester violence (column 5 in Table 2). The estimated effect, moreover, remains quantitatively stable and insignificant if I additionally consider other forms of militant behavior such as property destruction (column 5).

In summary, the baseline findings provide evidence for the hypothesis that a surge in (expected) mineral rents leads to a harsher and more violent crackdown on protesters but does not significantly affect vandalism or violence among protesters.

2.5 Robustness Checks

This subsection explores the sensitivity of the baseline results in several dimensions.

2.5.1 Omitted Variables

In the first set of robustness checks, I address concerns about potential omitted variables.

Neighborhood analysis Leaning on Berman et al. (2017), I implement a neighborhood fixed effects regression model. For this specification, I define the "control set" for each mining municipality as comprising the first- and second-degree neighboring *nonmining* municipalities. Each mining municipality and its neighbors form a "neighborhood group," with the price of the main mineral in the *mining* municipality also assigned to its neighbors. I include neighborhood $\times$ year fixed effects to absorb any constant and time-varying codeterminants of social conflict and mining activity across neighborhoods. Any identifying variation in this model derives from the differential reaction to price shocks

between mining municipalities and their inactive neighbors.

Table 3 presents the estimates for the neighborhood fixed effects model. The estimated probability of observing protest suppression by police forces is significantly higher in mining municipalities, while no significant difference for protester violence and riots is visible. The results are qualitatively stable when I consider the ten nearest neighbors by distance (Table B.3 in the Appendix).

Time-varying covariates In Table B.4 in the Appendix, I additionally control for municipality-specific factors varying at the monthly level that have been identified in the literature to be associated with conflict probability such as agricultural commodity prices (Dube and Vargas, 2013; **?**; McGuirk and Burke, 2020) and weather conditions (Hsiang et al., 2013; Harari and Ferrara, 2018). Movements in these variables could coincide with mineral price shocks and protest incidences. First, I introduce the main crop price in a municipality as an additional control. Following the coding of the main mineral price, the former is the agricultural commodity in a municipality with the highest production value. To calculate the total production value, I combine world commodity prices from the *World Bank* with crop-specific agricultural land cover circa 2000 from the *M3-Cropland* project (Monfreda et al., 2008).[19] The effect of crop prices is estimated to be indistinguishable from zero, while mineral prices retain their positive and significant coefficient. Similarly, the coefficient of interest remains statistically and quantitatively stable with the inclusion of temperature and rainfall as controls.

2.5.2 Measurement

Section B.3 in the Appendix shows that the baseline findings are robust to my considering alternative definitions and measures of *(i)* the municipality-specific mineral price, *(ii)* protest incidence, and *(iii)* mining activity.

Price index First, I investigate the robustness of the baseline results to the use of the price of all minerals mined in a municipality aggregated into a price index in lieu of

[19]The M3 crops data are also used by **?**, Harari and Ferrara (2018), and McGuirk and Burke (2020), among others.

the main mineral price. The price index is calculated as the weighted average price of all mined minerals, with the time-constant weights defined as the mineral's share of the total production value in a municipality over the sample period. The robust coefficient estimates for the alternative price measure are presented in Table B.5.

Outcomes Second, I check the sensitivity of the baseline estimates to the definition of protest incidence. In column 1 of Table B.6, I test if the likelihood of protests against mining projects overall increases. Again the estimates reveal no systemic relationship between mineral prices and protester behavior. Furthermore, estimates are robust to the the exclusion of months with protester riots (columns 2 to 5) or events with "unconfirmed" details—i.e., incidences where the number of injuries or arrests is not clearly stated (columns 6 and 7). In addition, Table B.7 in the Appendix presents estimates for the baseline specification when I use the location of the social conflict associated with a protest incident in lieu of the event location. That is, the protest incident is now coded as if it had happened in all municipalities spanned by the corresponding social conflict. While this alternative outcome definition leads to an (expected) loss of precision in the estimates, the magnitude of the effects remains quantitatively stable.

Production municipalities Third, I restrict the sample to municipalities with at least one year of mineral production. Municipalities with mining concessions but no production during the sample period are, hence, disregarded. Table B.8 in the Appendix shows that the focus on production municipalities results in positive but smaller and less precisely estimated coefficients; validating anecdotal evidence that social conflicts often emerge prior to the production stage when the chance to influence political and corporate decision is perceived as greater by local communities.

2.5.3 Econometric Specification

The baseline specification uses the the log of the level of the mineral price and does not consider potential temporal or spatial lags.

Level vs. first difference The use of price levels has evolved as the standard in the conflict literature (Berman et al., 2017, 2019) but requires that the price series be stationary. Unit root tests for each mineral-specific monthly price series (once purged from their common time components) suggest that price series are stationary, with the exception of monthly gold prices, for which the null hypothesis of the presence of a unit root marginally cannot be rejected at the 10% level (Figure B.1). To alleviate potential misspecification concerns, I present in Tables B.9 and B.10 in Section B.4 of the Appendix the estimates for the change in the natural logarithm of the main mineral price and the price index, respectively, from month $t-1$ to t. This focus on price growth results in qualitatively similar but less precisely estimated effects.

Temporal and spatial lags Further, allowing for spatial spillovers increases the impact of mineral prices (Table B.11), whereas including temporal lags B.12 has little influence on the cumulative effect. The results align with findings on spatial lags in the civil conflict literature (Berman et al., 2017; McGuirk and Burke, 2020; Berman et al., 2019).

2.5.4 Additional Robustness Checks

Seasonality Including month fixed effects in the baseline specification to account for potential seasonality of protests and violence leaves estimates quantitatively and qualitatively unchanged (Table B.13 in the Appendix).

World market share A potential concern with the proposed identification strategy is that social conflicts could impact mineral production to such an extent that they have some influence on world prices. Table B.2 in the Appendix presents in Panel A the world share of Peruvian mineral production disaggregated by mineral and year and in Panel B the corresponding maximum share among all municipalities. While Panel A confirms Peru's position as one of the main global mineral producers, Panel B illustrates that, with the exception of tin, no municipality has a world share above 5%, making any world prices movements for those minerals plausibly exogenous.[20] To check for the

[20]The San Rafael mine is located in the municipality of Antauta, in the province of Melgar, in the region of Puno.

possibility that local social conflicts affect world mineral pries, I drop all municipalities whose world market share ever exceeded 1% during the sample period from 2004 to 2019 for this analysis.[21] Table B.14 in the Appendix shows that the coefficients are virtually unchanged from the benchmark estimates, alleviating concerns about a potential violation of the exogeneity assumption.

Standard errors Next, I provide robustness results for different spatial kernels. In particular, I allow for unlimited serial correlation and spatial correlation within a radius of up to 50 km, 100 km, 250 km, 500 km, 750 km, and 1,000 km. For both dimensions, I assume a linear decay in distance. Table B.15 in the Appendix presents standard error estimates for these alternative levels of clustering. The estimates for protester arrests, injuries, and deaths retain significance at the 5% level regardless of which spatial radius is used, while the estimates for protester behavior remain indistinguishable from zero.

Multiple-hypothesis correction Finally, I address concerns that my baseline findings could suffer from an overrejection of the null hypotheses as a result of reuse of the identifying exogenous variation for multiple outcomes. I apply the *Romano–Wolf* multiple-hypothesis correction procedure, described in Romano and Wolf (2005a,b, 2016) and implemented by Clarke et al. (2020), to safeguard against false rejection of true null hypotheses. In recent work, Heath et al. (2022) show that the employed method performs well in a multitude of settings and across different dimensions. Table B.16 in the Appendix presents for each baseline outcome variable the original model p-value, the resampled p-value from 500 bootstraps, and the *Romano–Wolf* p-value corrected for multiple hypothesis testing.[22] The estimates for arrests, injuries, and the fatal use of force against protesters remain significant at the 10% level once multiple hypothesis testing is explicitly controlled for, while the null hypothesis of no effect of mineral rents on protester behavior cannot be rejected at common thresholds.

[21] Applying the 1% threshold is equivalent to excluding municipalities in the 90th percentile of world market share among all producing municipalities.

[22] A graphical illustration of the null distributions used to calculate the *Romano–Wolf* adjusted p-values for each of the five baseline outcome variables is given in Figure B.2 in the Appendix.

3 Mechanisms

3.1 Pro-Mining Local Politicians, State Repression, and Rent Seeking

In this section, I test the hypothesis that the election of a pro-mining as opposed to an anti-mining mayor increases state repression against protesters and corruption in the municipality.

3.1.1 Data

I combine data from various sources to build a municipality–candidate-level data set.

Data on election outcomes are obtained from the National Jury of Elections (JNE) online database *Infogob*. This includes the list of mayoral candidates and the results of municipal elections for 2002, 2006, 2010, 2014, and 2018. In addition, I obtain information on recall referenda for these five rounds of municipal elections to determine whether a sitting mayor was recalled prior to the end of their term.

Infogob also makes government plans ("*planes de gobierno*") of candidates available for each election round, with the exception of 2010.[23] These provide unique information on the political program and ideology of mayoral candidates in the context of highly fragmented and candidate-centered local level politics in Peru (Bland and Chirinos, 2014; Artiles et al., 2021).[24] The government plans were scraped from the *Infogob* website and are used to assemble a novel and comprehensive data set on the (stated) stance of candidates in regards to formal mining activity. Specifically, I classify candidates into either *anti-mining*, *pro-mining* or *neutral* toward formal mining based on their government plans published prior to the corresponding municipal election. The applied classification procedure can be summarized in the following steps. First, I preprocess the documents to obtain the raw text and use a semiautomatically derived mining keyword list to filter out passages concerned with formal mining. All government plans that do not contain any

[23] While links to government reports for elections in 2010 are provided on the government website, these links cannot be accessed, and the browser returns an error message (last accessed 20 August 2024).

[24] For instance, Artiles et al. (2021) report that, of the average 7.26 candidates running for mayor in 2014, only 36.9% ran for a national party.

mining keywords are classified as "unknown." Next, I use OpenAI's large language model *GPT-4* to answer five questions based on the passages extracted in the previous step to determine a candidate's sentiment toward formal mining. Note that a candidate's being classified as *anti-mining* does not require the candidate to oppose formal mining entirely; a focus on constraining current mining practices or on the negative environmental impacts of mining suffices. If the answers exhibit any inconsistencies, e.g., if the GPT-4 answers indicate a focus on both the positive and negative impacts of mining, I feed the text back to GPT-4 with a verification question to resolve the conflict.[25] Appendix Section A.2 provides a detailed discussion on the classification process.

Data on corruption is drawn from two sources. First, I use information from the *Encuesta Nacional de Hogares (ENAHO)* survey on local governance carried out yearly by Peru's national statistical agency since 2002. In particular, I make use of a consistently asked question to create a binary indicator on whether any member of the surveyed household was requested to, felt obligated to or voluntarily gave a bribe to a public official in the last 12 months, and zero otherwise.[26] Subsequently, I calculate for each year during a mayor's term the share of respondents in a municipality that experienced a bribery episode. Second, I retrieve data on the prosecution of corruption offenses in office by mayors (e.g., abuse of power, embezzlement, collusion, etc.) from the PPEDC's 2022 report on corruption in regional and local governments (Pacheco Palacios et al., 2022).[27] The report lists both, cases for which a former mayor has been sentenced and cases still at the investigative stage. The latter is of importance in light of the historically long judicial process associated with these crimes (Pacheco Palacios et al., 2022). The listed corruption cases are matched to elected mayors based on the accused mayor's name and the region where the crime took place in using the procedure described in Enamorado et al. (2019) and implemented in the R package `fastLink`.

[25] If the classification procedure returns an error at any of the stages, I resort to manual labeling of the government plan.

[26] For the baseline estimates, I exclude responses recorded in first year after the election, since any bribery related instances could be related to the former or current local administration. Results are, however, robust to the inclusion of this first year.

[27] In particular, I obtain information on the procedural stage of the case, the region where the (alleged) crime happened, the name of the accused, and the category of the crimes committed against public administration from Section 7 p. 209–613.

Finally, to compare how the election E of a mayor affected violence against activists and corruption in municipality i, I compute the difference between the average yearly level of outcome Y during a mayor's term and the level of Y in the year before the election (Marx et al., 2022):

$$\Delta Y_E = \underbrace{\frac{1}{k}\sum_{\tau=1}^{k} Y_{it_e+k}}_{\text{Post-election Average}} - \underbrace{Y_{it_e-1}}_{\text{Pre-election Value}},$$

where k denotes the number of years the mayor's term lasts.[28] Focusing on the change in levels rather than the level itself, allows me to control for differences across municipalities and time periods. For more details on the variables used in the analysis, please refer to Table A.1 in the Appendix.

3.1.2 Empirical Strategy

To test the hypothesis that the election of a pro-mining as opposed to an anti-mining mayor increases state repression against protesters and rent seeking, I restrict the set of elections to races with either a pro- or anti-mining winner and at least one anti-, respectively pro-mining opponent. I estimate the following Regression Discontinuity (RD) equation to address endogeneity concerns about unobserved variables correlated with the election of a pro-mining candidate and post-election outcomes:

$$\Delta Y_E = \delta_0 + \delta_1 T_E + \delta_2 X_E + \delta_3 T_i X_E + \epsilon_E, \tag{2}$$

where X_E, the running variable, denotes the difference between the vote share of the pro-mining candidate and their best-ranked anti-mining opponent and $T_E = \mathbb{1}[X_E > 0]$ if a pro-mining candidate wins the election. I use the the non-parametric method of Calonico et al. (2014) to estimate equation 2 and use their optimal MSE-bandwidth throughout with triangular kernel weights. Using this method, point estimates for (marginally) electing a pro-mining candidate over an anti-mining opponent (δ_1) and the robust p-value

[28] Note that the regular term length for a mayor is 4 years but the term may be cut short if the mayor is recalled.

adjusted for clustering at the regional level are reported.

For the coefficient of interest, δ_1, to be correctly identified two key assumptions have to hold: (1) there should be no manipulation around the cut-off, and (2) covariates potentially correlated with both the treatment and outcome variables should not significantly differ around the cut-off. To test for potential manipulation benefitting pro-mining candidates, I employ the density test suggested by Cattaneo et al. (2018). Figure C.1 in the Appendix reports this test and shows no discontinuous jump in the density of the running variable at cut-off (p-value = 0.464). Second, Table C.1 in the Appendix reports estimates using the same Regression Discontinuity (RD) equation as in 2 with predetermined covariates as outcomes. Reassuringly, election and municipality characteristics are found to not be significantly different on both sides of the threshold.

3.1.3 Results

We report the robust, bias-corrected estimates of δ_1, with additional specifications that linearly adjust for predetermined covariates, in Table 4. Figure C.2 in the Appendix provides a visual representation of the estimated relationship around the cut-off.

Electing a pro-mining candidate as opposed to an anti-mining candidate has a positive effect on both, the average prevalence of violent state repression and its frequency (columns 1-4). In contrast, the narrow election of a pro-mining mayor is not associated with more protester riots (Table C.2 in the Appendix),[29], suggesting—in line with the baseline findings—that the excessive use of force cannot be explained by changes in protester behavior. Columns 5-6 of Table 4 present estimates for the change in reported bribery instances by citizens once a pro-mining mayor took office. While the sample size is small, I find a positive and precisely estimated effect on corruption once I adjust for predetermined covariates to improve precision (Cattaneo et al., 2020). Reassuringly, the number of respondents does not vary around the cut-off (Table C.1), alleviating concerns about potential underreporting in bribery-related questions after a pro-mining mayor takes office due to, e.g., more non-responses out of fear of repercussions.[30] In contrast,

[29]Figure C.3 in the Appendix provides a visual representation of the estimated treatment effects.

[30]Herrera et al. (2007) also show that the non-response rate of the ENAHO governance module is

I find an insignificant negative effect of narrowly electing a pro-mining candidate on the probability that the mayor will be investigated for crimes against public administration during their tenure (columns 7-8). The apparent dichotomy between the reported increase in corruption and the probability of investigation into corruption offenses indicates political capture of the judicial process. Furthermore, Table C.3 in the Appendix shows that pro-mining mayors are more likely to face initial investigations than anti-mining mayors but are less likely to face formal charges;[31] a pattern consistent with pro-mining mayors being more effective in stalling initial investigations against them.

Taken together, the RD results provide support for the notion that the estimated positive relationship between mineral rents and violent state repression in Section 2.4 is connected to rent seeking activities of local authorities.

3.2 Democratic Accountability

In this subsection, I examine if mayors are held politically accountable for police violence during their term.

Empirical Strategy To test this channel, I again implement a continuity-based Regression Discontinuity Design (RDD). In secific, I comparing differences in the reelection outcomes in $t+1$ of the top two candidates in the municipal election at time t::

$$Y_{jit+1} = \delta_0 + \delta_1 \text{Incumbent}_{jit} + \delta_2 \text{Incumbent}_{jit} A_{jit} + \delta_3 A_{jit} + \delta_4 \text{Incumbent}_{jit} X_{jit} + \delta_5 \text{A}_{jit} X_{jit} + \delta_6 \text{Incumbent}_{jit} A_{jit} X_{jit} + \epsilon_{jit}, \quad (3)$$

where j denotes a candidate running in municipality i at time t; X_{jit} is the running variable denoting the difference in the vote share between the winner of the municipal election i and the runner-up at time t; Y_{jit+1} is a measure of j's electoral performance in the subsequent election; $\text{Incumbent}_{jjt} = \mathbb{1}[X_{jit} > 0]$ defines the individual-level treatment (winning municipal election i at time t) and A_{ijt} is a binary indicator equaling one if

significantly lower than the overall survey non-response rate.

[31] Figure C.4 in the Appendix provides a visual representation of the estimated treatment effects.

the elected candidate's term was characterized by at least one incident of violent state repression, and zero otherwise.

If incumbents and runner-ups are indistinguishable at the threshold except for their incumbency status, δ_1 estimates the causal effect of incumbency on subsequent electoral outcomes at the cutoff, whereas δ_2 gauges the effect of police violence on electoral outcomes. If mayors are held accountable for state violence against their constituents, the incumbency effect is expected to be smaller ($\delta_2 < 0$).

I distinguish between *unconditional* and *conditional* effect estimates to account for selection into rerunning (De Magalhaes, 2015); a decision unlikely to be random, with only approximately 36% of incumbents running in $t+1$ as opposed to 54% of runners-up. First, I consider the *unconditional* estimate on the reelection status and vote share of the top-two candidates in period $t+1$, regardless of whether they actually run in $t+1$. To interpret the *unconditional* estimates, I assess differences in the probability of rerunning before estimating the *conditional* effect for the restricted sample of candidates who run again in the next election. These outcome variables are available for municipal elections in 2002, 2006, 2010, after congress banned reelection in 2015.[32] Table A.1 in the Appendix provides more details on independent and dependent variables used in this analysis.

I estimate equation 3 using the non-parametric and robust regression discontinuity (RD) estimator of Calonico et al. (2014) with MSE-optimal bandwidth and triangular weights. Robust standard errors are clustered at the regional level. By construction, the symmetric structure of the estimation, with two observations per municipal election, implies that the sample is perfectly balanced for all municipality-level variables and the density of the victory margin is identical on both sides of the cutoff.

Results Table 5 presents estimates of δ_1 and δ_2 from equation 3. Two main takeaways emerge. First, and in line with previous studies of local elections in Peru and other democracies with a highly fragmented, weak party system (e.g. Klašnja and Titiunik, 2017; Holland and Incio, 2019; Weaver, 2021), I find a significant incumbent disadvan-

[32]Note that the data on police violence is incomplete for mayoral terms between 2002-2006, since information on this variable is only available from March 2004 onward.

tage. Candidates who narrowly won in t are unconditionally less likely to win in $t+1$ (column 1) and garner significantly less votes in $t+1$ (column 2) than runner-ups who narrowly lost. While part of the *unconditional* effect can be explained by the significantly lower probability of re-entry of winners compared to runner-ups (column 3), the incumbency disadvantage persists even when conditioning on running in $t+1$ (columns 4 to 6). Second, both, the *unconditional* and *conditional* incumbency disadvantage is—if anything—smaller for incumbents whose term was characterized by violent state repression. This pattern is qualitatively robust to estimating two separate RD equations for mayoral terms *with* and *without* police violence (Table D.1).

Overall, the results suggest that local democratic institutions are ineffective in holding authorities accountable for the violent state repression of protests.

3.3 Is Violent State Repression Effective?

In this section, I take advantage of a unique feature of the social conflict data set: each social conflict and associated events are tracked over time. This allows me to investigate whether the violent repression of social protests proves effective (from the perspective of the repressing authorities). In particular, I look at two dimensions: *(i)* the duration of the social conflict and *(ii)* the final outcome of the social conflict. For brevity, results for the first dimension are presented in Appendix Section E.2. In summary, the survival analysis finds that violence against protesters prolongs the duration of social conflicts. Conditional on conflict duration, local authorities may nevertheless achieve their preferred outcome—the subject of study for the remainder of this section.

3.3.1 Data

The end of social conflicts can be broadly categorized into three different outcomes: *(1)* signing of a resolution agreement *(2)* "removal" of the conflict due to inactivity, and *(3)* right-censoring of the conflict due to the end of the study period or loss to follow-up.[33] A *resolution* agreement in the view of mining corporations and local authorities

[33]Only in 4 of the 76 instances where a social conflict was right-censored, the conflict was right-censored due to loss to follow-up, i.e. because the Peruvian Office of the Ombudsman discontinued their

is the "unwanted" outcome: it usually contains some form of concession to the activists such as compensation payments for the (illegal) use of communal land or damages to the environment. In contrast, the *removal* of a conflict requires that the social conflict has been inactive for such a prolonged period of time that the Peruvian Office of the Ombudsman views the conflict has having ended and discontinues its reporting. This outcome can, hence, be viewed as the preferred or "sought-after" one.

I define the end of the conflict as the last month when the social conflict was still active. This definition allows me to account for two aspects of the data. First, the Peruvian Ombudsman does not have a fixed threshold for time of inactivity beyond which social conflicts are no longer tracked; i.e., conflicts can be removed after 8 months of inactivity or two years. Second, some social conflicts become inactive after the successful negotiation of a resolution agreement but before its signing.

I combine the data on social conflicts with an extensive set of time-varying and time-invariant conflict-specific covariates. To account for the sparsity of police violence events over the course of social conflicts, the unit of observation for the subsequent analysis is the conflict-year level. The set of time-varying controls comprises past incidents of police violence, protests and protester riots, the change in the main mineral price, *canon minero* and royalty payments to the municipalities where the conflict is situated, and nighttime luminosity as a proxy for economic development (e.g. Henderson et al., 2012). Time-invariant *baseline* covariates include affected municipality characteristics such as a binary indicator for the presence of indigenous communities or native lands, population density, river density, road density, area covered by lakes, total surface area, and mean elevation. The additional baseline covariates provide information on the characteristics of the corporate owner of the contentious mining project and encompass the market value of the majority owner as well as a binary indicator for foreign ownership. Table A.1 in the Appendix provides a comprehensive list of all variables used in this section and Table E.1 presents descriptive statistics.

reporting on the conflict without explanation. The remainder of conflicts was censored due to the sample period ending in December 2019.

3.3.2 Empirical Strategy

In the context of social conflicts, the decision of local authorities on whether to crack down on protesters in period t to reach the sought-after conflict outcome Y is a dynamic process that takes into account past developments. Traditional regression models fail to appropriately adjust for time-dependent observed confounders that are affected by previous treatments (s. Appendix Appendix E.3 for more details). In the current setting, we additionally face the challenge of not observing the final outcome for all social conflicts. The literature has so far primarily relied on (semi-)parametric marginal structural models (MSM) to adjust for time-dependent observed confounders while avoiding posttreatment bias (e.g. Blackwell, 2013; Ladam et al., 2018). The estimation of inverse probability weights, however, makes these models highly sensitive to misspecification of the treatment assignment model. Several extensions have since been proposed to make estimates more robust to misspecification (e.g. Imai and Ratkovic, 2015; Zhou and Wodtke, 2020) but two important issues remain that are particularly relevant in the context of this study. First, an identifying assumption of MSM is that each subject in the population has a non-zero chance of receiving all possible treatment sequences.[34] This so called "positvity" assumption is unlikely to be fulfilled in the setting under investigation with police violence being observed only in 8.4% of all conflict years.[35] Second, even if *positivity* technically holds, with only a small fraction of the possible treatment sequences observed in the sample, propensity scores will be extreme (i.e. close to zero or one) for some conflicts, leading to unstable weight and treatment effect estimates.

Incremental causal effects To recover a causal estimand in the context of this study, I therefore rely on a non-parametric estimator based on *incremental propensity scores interventions* proposed by Kennedy (2019) and extended in Kim et al. (2021) that is robust to violations (or near-violations) of the *positivity* assumption and accommodates time-varying censoring. In contrast to other estimators such as MSM, the aim of this

[34]This assumption is similar to the assumption of *common support* in the matching literature.

[35]In detail, in only 75 of the 896 conflict-years in the sample at least one incidence of police violence against protesters is observed ($A_t = 1$).

estimator is to identify the causal effect of increasing or decreasing everyone's odds of receiving treatment rather than the effect of receiving treatment or not. More formally, the observed propensity score $\pi_t(\mathbf{h}_t) = \mathbb{P}(A_t = 1 \mid \mathbf{H}_t, R_{t-1} = 1)$, is multiplied by a *user-specified* increment parameter $\delta \in (0, \infty)$ to obtain the intervention distribution:

$$q_t(\mathbf{h}_t, \delta, \pi_t) = \frac{\delta \pi_t(\mathbf{h}_t)}{\delta \pi_t(\mathbf{h}_t) + 1 - \pi_t(\mathbf{h}_t)}, \tag{4}$$

where $\mathbf{H}_t = \left(\overline{\mathbf{X}}_t, \overline{A}_{t-1}\right)$ denotes all covariate $\overline{\mathbf{X}}_t = (\mathbf{X}_1, ..., \mathbf{X}_t))$ and treatment $\overline{\mathbf{A}}_{t-1} = (\mathbf{A}_1, ..., \mathbf{A}_{t-1})$ history up to just prior to treatment at time t, and R_{t-1} is a binary indicator equal to one if the conflict is still in the sample at $t-1$ (is not right-censored by $t-1$). Rewriting equation 4 unveils that δ has an intuitive interpretation as the odds ratio comparing the odds of q_t to the odds of π_t:

$$\delta = \frac{q_t / (1 - q_t)}{\pi_t / (1 - \pi_t)}.$$

If $\delta > 1$ ($\delta < 1$) the intervention increases (decreases) the odds of receiving treatment, while the treatment process is left unchanged if $\delta = 1$ ($q_t = \pi_t$). The causal estimand of interest is the average counterfactual outcome (across social conflicts) if the odds of treatment were multiplied by δ, i.e. $\psi(\delta) = \mathbb{E}\left[Y^{\mathbf{Q}(\delta)}\right]$, where $Y^{\mathbf{Q}(\delta)}$ is the potential outcome under the treatment process $(q_1, ..., q_T)$.

Identification and Estimation Kennedy (2019) shows that the incremental effect $\psi(\delta)$ is correctly identified under (A1) *consistency*, and (A2) *exchangability*.[36] *Consistency* requires that the observed outcomes equal the corresponding potential outcomes under the observed treatment process. This assumption would be violated if, for instance, police violence in one conflict has an effect on the final outcome of another social conflict. *Exchangability* assumes that treatment each period is as good as random conditional on the past. In other words, *exchangability* demands that there are no unmeasured confounders. The extensive set of time-varying and baseline covariates in conjunction with

[36]Note that the *exchangability* assumption is also referred to as the *sequential ignorability* assumption in the literature (e.g. Blackwell, 2013; Imai and Ratkovic, 2015).

the use of non-parametric machine learning methods that are effective in detecting nonlinearities and interactions lends credibility to the *exchangability* assumption (Montgomery and Olivella, 2018). In the case of right-censoring, two additional assumption have to hold (Kim et al., 2021): (A3) dropout conditional on the past is as good as random and (A4) each subject has a non-zero chance of staying in the study at the next timepoint. Assumptions that are likely to be fulfilled in the context of this study with right-censoring being—with the exception of four conflicts that were dropped from the watchlist of the Peruvian Ombudsman without explanation—the result of the investigation period ending in December 2019 due to the start of the Covid-19 pandemic in 2020.

The estimation of the incremental effect $\psi(\delta)$ for each δ involves the following steps:[37]

1. *Sample Splitting.* Randomly split the full sample into k samples stratified by treatment.[38] For a given k (here $k = 2$), define testing (including all conflicts selected into split k) and training (including all conflicts not selected into split k) data sets. The sample splitting allows the use of *double machine learning* (Chernozhukov et al., 2018) which yields asymptotically normal and efficient estimators, all while placing no restrictions on the complexity of the nuisance estimators (Kennedy, 2019).

2. *Estimate nuisance parameters.* For each conflict and timepoint, regress the treatment indicator on historical variables ($A_t \sim \mathbf{H}_t$) and the indicator for not dropping out on current treatment status and past values ($R_t \sim A_t + \mathbf{H}_t$) within the *training data* and use the fitted models to predict $\hat{\pi}_t$ and $\hat{\omega}_t$ (i.e. the probability of not dropping out) in the *full sample*. Subsequently, the predicted values and the observed treatment process are used to build cumulative weights:

$$W_t = R_t \prod_{s=1}^{t} \frac{\delta A_s + 1 - A_s}{\delta \hat{\pi}_s + 1 - \hat{\pi}_s} \cdot \frac{1}{\hat{\omega}_s}$$

3. *Estimate the pseudo regression functions.* For each timepoint $t = T, T-1,...,1$ (starting with $M_{t+1} = Y$) regress M_{t+1} on treatment and the past ($M_{t+1} \sim A_t + \mathbf{H}_t$)

[37] For brevity, I provide a stylized description of the underlying algorithm and refer for a more detailed description to Kim et al. (2021).

[38] Given the sparsity of police violence in the sample, stratified sampling ensures that the training data has enough positive example for the machine algorithm to learn the model parameters.

within the *training* data set. The model output is then used to predict the outcomes $m_t(\mathbf{H}_t, a_t = 1)$ and $m_t(\mathbf{H}_t, a_t = 0)$ within the conflicts that have not dropped out at this point. Use these predicted outcomes $\hat{m}_t$ together with A_t, $\hat{\pi}_t$, R_t, $\hat{\omega}_t$, and δ to compute the *pseudo-outcome* M_t for the *full sample.*

4. *Estimate the incremental effect* Combine the results from the steps above to compute $\varphi = W_T Y + \sum_{t=1}^{T} W_t M_t$ for all conflicts in the *testing data* and let $\hat{\psi}^k(\delta)$ be its average. Repeat the steps above for the other sample splits, and obtain the final estimate for the incremental effect $\hat{\psi}(\delta)$ by averaging $\hat{\psi}^k(\delta)$ across sample splits.

5. *Inference.* Pointwise confidence intervals for $\hat{\psi}(\delta)$ can be constructed by using the empirical variance of the estimated influence functions φ (s. Kennedy, 2019; Kim et al., 2021). In addition, uniform confidence intervals can be obtained using a multiplier bootstrap (s. Kennedy, 2019) that allow to draw inference over the whole range of δ values.

3.3.3 Results

Figure 3 presents the incremental effect curve $\hat{\psi}(\delta)$ reporting the estimated proportion of conflicts ending in a resolution agreement if the odds of violent state repression are multiplied by $\delta \in [0.1, 3.0]$. All nuisance functions π_t, ω_t and m_t are estimated using the *superleaner* ensemble algorithm (van der Laan et al., 2007) implemented in the `SuperLearner` R package (Polley et al., 2024). The *superleaner* uses cross-validation to compare the performance of each ensemble method and returns the optimal weighted average of those models using the test data performance.[39] Specifically, I use five-fold cross-validation and an ensemble of random forest via `ranger` (Wright and Ziegler, 2017), extreme gradient boosting via `xgboost` (Chen and Guestrin, 2016), and k-nearest neighbor regression via `kernelKnn` (Mouselimis, 2023). 90% pointwise and uniform confidence intervals are calculated using $10,000$ bootstrap replications.

A steady pattern emerges for the relationship between the proportion of social conflicts

[39] Polley and Van der Laan (2010) show that the *superleaner* ensemble algorithm is asymptotically as accurate as the best possible prediction algorithm in the ensemble that is tested.

ending with an undesired resolution agreement and state repression: decreasing (increasing) the odds of police violence against protesters increases (decreases) the prevalence of resolution agreements. In particular, with no intervention on the treatment process ($\delta = 1$) but an intervention to remove censoring, the proportion of conflicts ending with a resolution agreement is estimated to be 59.60% (90% CI = 52.27%, 66.93%) and very close to the observed resolution prevalence of 61.06%. Decreasing the odds of police violence proportionally for all conflicts ($\delta < 1$) leads to a steady increase in resolution agreement prevalence. For instance, for $\delta = 0.3$ the estimated proportion of resolution agreements is 69.86% (90% CI = 59.22%, 80.51%). This implies that compared to no intervention ($\delta = 1$), the resolution prevalence increases by 10.26% (90% CI = 3.01%, 17.52%).[40] Conversely, increasing the odds of violent state repression ($\delta > 1$) steadily decreases the incidence of resolution agreements. Effects for large δ are, however, imprecisely estimated, mainly due to few conflicts in the sample with multiple years of police violence. Broad confidence bands for large δ values also do not allow to reject the null hypothesis of no effect across the full δ range from 0.1 to 3.0 at conventional confidence levels (the uniform confidence band contains a horizontal line).

Overall, the steady pattern uncovered in Figure 3 provides a some first empirical evidence for the effectiveness of violent state repression in reaching the social conflict outcome sought-after by local authorities.

4 Conclusion

This paper introduces a new data set on violent suppression of socioenvironmental advocacy against formal mining activity in Peru. Leveraging the spatial and temporal granularity of the data and exogenous changes in world mineral prices, I provide causal evidence on the positive relationship between local mineral rents and violent confrontations between protesters opposing local mining projects and police forces. Quantitatively, I estimate that a one-percentage-point increase in world prices of the main mineral mined

[40] Resolution agreement incidence estimates and estimated differences compared to no intervention for additional δ values are provided in Table E.3 in the Appendix.

in the municipality more than doubles the probability to observe use of excessive force against protesters relative to the mean. Further, I show that this escalation in social conflicts is independent of militant protester behavior such as rioting or violence against officers.

These baseline results are complemented with empirical evidence from municipality elections. Using a regression discontinuity design, I find that the narrow election of a *pro-mining* as opposed to an *anti-mining* candidate increases the probability of police violence against activists. Moreover, empirical results suggest the political capture of legal and democratic institutions. In particular, I find that the election of a *pro-mining* candidate increases the prevalence of corruption as experienced by constituents but this expansion is not reflected in a coincident increase in investigations against *pro-mining* mayors for corruption offenses in office. Further, I show that mayors whose term was characterized by police violence against protesters were not punished at the ballot box in the next election, suggesting that democratic accountability is compromised in local elections in Peru.

Finally, I take a first step in the direction of providing causal estimates on the efficacy of violent state repression in diffusing protest movements. Taking advantage of unique information on the full timeline of each social conflict in the data set, I show that the strategic use of violence against protesters affects the final outcome of the conflict. Formal resolution agreements, which commonly involves financial compensation for environmental damages or use of land for local communities, are estimated to become more (less) prevalent if the odds of police violence against protesters are decreased (increased).

This study contributes to the emerging literature on the political economy of human rights. Its findings highlight that violence against agents of civil society constitutes an important negative externality of local mining activity that is not restricted to authoritarian regimes. Rather, it highlights the potential adverse effects of central government transfers of natural resource rents to local governments even if the use of those funds is regulated by law and free elections should ensure democratic accountability of local authorities. This study's findings can, therefore, serve as an illustrative precedent to

policymakers for the importance of designing redistribution schemes that internalize negative externalities and align incentives of corporations, authorities, and local community activists.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D. and Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity and poverty*. Profile, London.

Acemoglu, D. and Robinson, J. A. (2020). *The narrow corridor: States, societies, and the fate of liberty*. Penguin.

Agüero, J. M., Balcázar, C. F., Maldonado, S., and Ñopo, H. (2021). The value of redistribution: Natural resources and the formation of human capital under weak institutions. *Journal of Development Economics*, 148:102581.

Aragón, F. M. and Rud, J. P. (2013). Natural resources and local communities: Evidence from a peruvian gold mine. *American Economic Journal: Economic Policy*, 5(2):1–25.

Arellano-Yanguas, J. (2011). Aggravating the resource curse: Decentralisation, mining and conflict in peru. *Journal of Development Studies*, 47(4):617–638.

Aresti, M. L. (2016). Mineral revenue sharing in peru. *Natural Resource Governance Institute*, pages 18–19.

Artiles, M., Kleine-Rueschkamp, L., and León-Ciliotta, G. (2021). Accountability, Political Capture, and Selection Into Politics: Evidence from Peruvian Municipalities. *Review of Economics and Statistics*, 103(2):397–411.

Asher, S. and Novosad, P. (2023). Rent-Seeking and criminal politicians: Evidence from mining booms. *Review of Economics and Statistics*, 105(1):20–39.

Baragwanath Vogel, K. (2021). *The effect of oil windfalls on political corruption: Evidence from Brazil*. PhD thesis, UC San Diego.

Bazzi, S. and Blattman, C. (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4):1–38.

Berman, N. and Couttenier, M. (2015). External shocks, internal shots: The geography of civil conflicts. *Review of Economics and Statistics*, 97(4):758–776.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review*, 107(6):1564–1610.

Berman, N., Couttenier, M., and Soubeyran, R. (2019). Fertile ground for conflict. *Journal of the European Economic Association*, 19(1):82–127.

Besley, T. and Persson, T. (2011). The logic of political violence. *Quarterly Journal of Economics*, 126(3):1411–1445.

Blackwell, M. (2013). A framework for dynamic causal inference in political science. *American Journal of Political Science*, 57(2):504–520.

Blair, G., Christensen, D., and Rudkin, A. (2021). Do commodity price shocks cause armed conflict? a meta-analysis of natural experiments. *American Political Science Review*, 115(2):709–716.

Bland, G. and Chirinos, L. A. (2014). Democratization through contention? regional and local governance conflict in peru. *Latin American Politics and Society*, 56(1):73–97.

Bodory, H., Huber, M., and Lafférs, L. (2022). Evaluating (weighted) dynamic treatment effects by double machine learning. *The Econometrics Journal*, 25(3):628–648.

Bonvini, M., McClean, A., Branson, Z., and Kennedy, E. H. (2023). Incremental causal effects: An introduction and review. In *Handbook of Matching and Weighting Adjustments for Causal Inference*, volume 1, pages 349–372. CRC Press, 1 edition.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *American Economic Review*, 103(5):1759–96.

Butt, N., Lambrick, F., Menton, M., and Renwick, A. (2019). The supply chain of violence. *Nature Sustainability*, 2(8):742–747.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Canavire-Bacarreza, G. J., Martínez-Vázquez, J., and Sepulveda, C. F. (2012). Subnational revenue mobilization in peru. *Andrew Young School of Policy Studies Research Paper Series*, (12-22):12–09.

Carreri, M. and Dube, O. (2017). Do natural resources influence who comes to power, and how? *The Journal of Politics*, 79(2):502–518.

Castellares, R., Fouché, M., et al. (2017). The determinants of social conflicts in mining production areas. Technical report.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2020). *A practical introduction to regression discontinuity designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, 18(1):234–261.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA. Association for Computing Machinery.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Clarke, D., Romano, J. P., and Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in Stata. *Stata Journal*, 20(4):812–843.

Colella, F., Lalive, R., Sakalli, S. O., and Thoenig, M. (2023). acreg: Arbitrary correlation regression. *Stata Journal*, 23(1):119–147.

Crabtree, J. (2014). Funding local government: Use (and abuse) of Peru's Canon system. *Bulletin of Latin American Research*, 33(4):452–467.

Cui, Y., Kosorok, M. R., Sverdrup, E., Wager, S., and Zhu, R. (2023). Estimating heterogeneous treatment effects with right-censored data via causal survival forests. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(2):179–211.

Davies, S., Pettersson, T., and Öberg, M. (2022). Organized violence 1989–2021 and drone warfare. *Journal of Peace Research*, 59(4):593–610.

de Benedictis-Kessner, J. (2018). Off-cycle and out of office: Election timing and the incumbency advantage. *Journal of Politics*, 80(1):119–132.

De Magalhaes, L. (2015). Incumbency effects in a comparative perspective: Evidence from Brazilian mayoral elections. *Political Analysis*, 23(1):113–126.

Dube, O. and Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from colombia. *Review of Economic Studies*, 80(4 (285)):1384–1421.

Enamorado, T., Fifield, B., and Imai, K. (2019). Using a probabilistic model to assist merging of large-scale administrative records. *American Political Science Review*, 113(2):353–371.

Fetzer, T. and Kyburz, S. (2024). Cohesive institutions and political violence. *Review of Economics and Statistics*, 106(1):133–150.

Fetzer, T. and Marden, S. (2017). Take what you can: Property rights, contestability and conflict. *Economic Journal*, 127(601):757–783.

Grasse, D. (2022). Oil crops and social conflict: Evidence from indonesia. *Journal of Conflict Resolution*, 66(7-8):1422–1448.

Harari, M. and Ferrara, E. L. (2018). Conflict, climate, and cells: A disaggregated analysis. *Review of Economics and Statistics*, 100(4):594–608.

Haslam, P. A. and Ary Tanimoune, N. (2016). The determinants of social conflict in the latin american mining sector: New evidence with quantitative data. *World Development*, 78:401–419.

Heath, D., Ringgenberg, M. C., Samadi, M., and Werner, I. M. (2022). Reusing natural experiments. *Journal of Finance*, Forthcoming.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Herrera, J., Razafindrakoto, M., and Roubaud, F. (2007). Governance, democracy and poverty reduction: Lessons drawn from household surveys in sub-saharan africa and latin america. *International Statistical Review*, 75(1):70–95.

Holland, A. C. and Incio, J. (2019). Imperfect recall: The politics of subnational office removals. *Comparative Political Studies*, 52(5):777–805.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, 341(6151):1235367.

Imai, K. and Ratkovic, M. (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association*, 110(511):1013–1023.

Kennedy, E. H. (2019). Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, 114(526):645–656.

Kim, K., Kennedy, E. H., and Naimi, A. I. (2021). Incremental intervention effects in studies with dropout and many timepoints. *Journal of Causal Inference*, 9(1):302–344.

Klašnja, M. and Titiunik, R. (2017). The incumbency curse: Weak parties, term limits, and unfulfilled accountability. *American Political Science Review*, 111(1):129–148.

Ladam, C., Harden, J. J., and Windett, J. H. (2018). Prominent role models: High-profile female politicians and the emergence of women as candidates for public office. *American Journal of Political Science*, 62(2):369–381.

Lewis, B. D., Nguyen, H. T., and Hendrawan, A. (2020). Political accountability and public service delivery in decentralized Indonesia: Incumbency advantage and the performance of second term mayors. *European Journal of Political Economy*, 64:101910.

Li, X., Zhou, Y., Zhao, M., and Zhao, X. (2020). A harmonized global nighttime light dataset 1992–2018. *Scientific data*, 7(1):168.

Limodio, N. (2022). Terrorism financing, recruitment, and attacks. *Econometrica*, 90(4):1711–1742.

Loayza, N. and Rigolini, J. (2016). The local impact of mining on poverty and inequality: Evidence from the commodity boom in peru. *World Development*, 84:219–234.

Maldonado, S. (2011). Resource windfall and corruption: Evidence from peru. *Unpublished manuscript. UC Berkeley.*

Maldonado, S. (2017). The non-monotonic political effects of resource booms. *SSRN Working Paper 3153301.*

Maldonado, S. and Ardanaz, M. (2022). Natural resource windfalls and efficiency in local government expenditure: Evidence from peru. *Economics and Politics.*

Marshall, M. G., Gurr, T. R., Davenport, C., and Jaggers, K. (2002). Polity iv, 1800-1999: Comments on munck and verkuilen. *Comparative Political Studies*, 35(1):40–45.

Marx, B., Pons, V., and Rollet, V. (2022). Electoral turnovers. Working Paper 29766, National Bureau of Economic Research.

McGuirk, E. and Burke, M. (2020). The economic origins of conflict in africa. *Journal of Political Economy*, 128(10):3940–3997.

Monfreda, C., Ramankutty, N., and Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22(1).

Montgomery, J. M. and Olivella, S. (2018). Tree-based models for political science data. *American Journal of Political Science*, 62(3):729–744.

Mouselimis, L. (2023). *KernelKnn: Kernel k Nearest Neighbors*. R package version 1.1.5.

Orihuela, J. C., Pérez, C. A., and Huaroto, C. (2019). Do fiscal windfalls increase mining conflicts? not always. *Extractive Industries and Society*, 6(2):313–318.

Pacheco Palacios, J. A., Yudith, V. E., and Cesar Augusto, O. R. (2022). El avance de la corrupción desde la perspectiva de la defensa jurídica del estado: Gobiernos regionales y locales. *Procuraduría Pública Especializada en Delitos de Corrupción*.

Pique, R. (2019). Higher pay, worse outcomes? the impact of mayoral wages on local government quality in peru. *Journal of Public Economics*, 173:1–20.

Polley, E., LeDell, E., Kennedy, C., and van der Laan, M. (2024). *SuperLearner: Super Learner Prediction*. R package version 2.0-29.

Polley, E. C. and Van der Laan, M. J. (2010). Super learner in prediction. *U.C. Berkeley Division of Biostatistics Working Paper Series*, Working Paper 266.

Romano, J. P. and Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108.

Romano, J. P. and Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

Romano, J. P. and Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, 113:38–40.

Sánchez De La Sierra, R. (2020). On the origins of the state: Stationary bandits and taxation in eastern congo. *Journal of Political Economy*, 128(1):32–74.

Sant'Anna, P. H. (2016). Program evaluation with right-censored data. *arXiv preprint arXiv:1604.02642*.

Sexton, R. (2020). Unpacking the local resource curse: How externalities and governance shape social conflict. *Journal of Conflict Resolution*, 64(4):640–673.

Sundberg, R. and Melander, E. (2013). Introducing the ucdp georeferenced event dataset. *Journal of Peace Research*, 50(4):523–532.

Tian, L., Zhao, L., and Wei, L. J. (2013). Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. *Biostatistics*, 15(2):222–233.

Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., Schrag, D., Takeuchi, M., Uyama, Y., Zhao, L., Skali, H., Solomon, S., Jacobus, S., Hughes, M., Packer, M., and Wei, L.-J. (2014). Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology*, 32(22):2380–2385.

van der Laan, M. J., Polley, E. C., and Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1).

Weaver, J. A. (2021). Electoral dis-connection: The limits of reelection in contexts of weak accountability. *The Journal of Politics*, 83(4):1462–1477.

Wright, M. N. and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77(1):1–17.

Zhou, X. and Wodtke, G. T. (2020). Residual balancing: A method of constructing weights for marginal structural models. *Political analysis*, 28(4):487–506.

Zingales, L. (2017). Towards a political theory of the firm. *Journal of Economic Perspectives*, 31(3):113–30.

5 Tables

Table 1: Descriptive Statistics: Municipality Level (in %)

		Force used against protesters			Protester behavior	
		Pr(Arrests > 0)	Pr(Injuries > 0)	Pr(Casualties > 0)	Pr(Violence > 0)	Pr(Riots > 0)
All	N	355870	355870	355870	355870	355870
	Mean	0.013	0.038	0.008	0.008	0.012
	SD	1.161	1.955	0.871	0.918	1.073
Production	N	40850	40850	40850	40850	40850
	Mean	0.049	0.149	0.039	0.022	0.037
	SD	2.212	3.861	1.979	1.484	1.916
Concessions	N	53390	53390	53390	53390	53390
	Mean	0.036	0.079	0.011	0.024	0.030
	SD	1.886	2.804	1.060	1.560	1.731
None	N	261630	261630	261630	261630	261630
	Mean	0.003	0.013	0.002	0.003	0.004
	SD	0.587	1.123	0.437	0.553	0.618

Notes: All variables are presented in percent. Author's computation on basis of MIMEM and SNL Minings & Metals database.

Table 2: Social Conflict and Mineral Prices

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price)	0.0004**	0.0019***	0.0007**	0.0002	0.0002
	(0.0002)	(0.0007)	(0.0003)	(0.0002)	(0.0002)
Dep. variable mean	0.0004	0.0011	0.0002	0.0002	0.0003
Municip. $\times$ year FEs	✓	✓	✓	✓	✓
Observations	94241	94241	94241	94241	94241

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Neighborhood Analysis

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
ln(Price)	0.0001	-0.0003	-0.0001	-0.0001	0.0000
	(0.0001)	(0.0003)	(0.0001)	(0.0002)	(0.0002)
$M\times$ ln(Price)	0.0010*	0.0022**	0.0007*	0.0005	0.0004
	(0.0006)	(0.0009)	(0.0004)	(0.0004)	(0.0005)
Dep. variable mean	0.00009	0.00028	0.00004	0.00008	0.0001
Neighbor $\times$ year FEs	✓	✓	✓	✓	✓
Month FEs	✓	✓	✓	✓	✓
Observations	1149880	1149880	1149880	1149880	1149880

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Robust standard errors are clustered at the neighbor and month level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: The Effect of Electing a Pro-mining Mayor on State Repression and Corruption

	Police Violence Prevalence		Police Violence Frequency		Reported Corruption		Corruption Investigations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pro-Mining Mayor	0.081	0.097*	0.007	0.008*	0.039	0.033***	-0.096	-0.062
Cluster-robust p-value	0.108	0.084	0.11	0.086	0.488	0.007	0.761	0.832
90% CI	[-0.002, 0.192]	[0.005, 0.196]	[-0.002, 0.017]	[-0.001, 0.017]	[-0.08, 0.167]	[0.021, 0.134]	[-0.543, 0.397]	[-0.478, 0.385]
Baseline Covariates		✓		✓		✓		✓
Bandwidth	14.2	14.4	13.4	12.8	10.5	9.9	13.0	13.6
Dep. variable mean	0.010	0.010	0.001	0.001	-0.003	-0.003	0.167	0.167
Observations	230	230	230	230	54	54	234	230
Effective obs.	115	120	111	105	21	20	108	111

Notes: Local linear regression estimates from Calonico et al. (2014) with triangular kernel weights and optimal MSE bandwidth are reported. Even columns present local covariate-adjusted linear estimates. 90% robust confidence intervals and p-values adjusted for clustering at the regional level, the number of observations in the sample and in the bandwidth, the optimal MSE bandwidth, and the dependent variable mean are reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

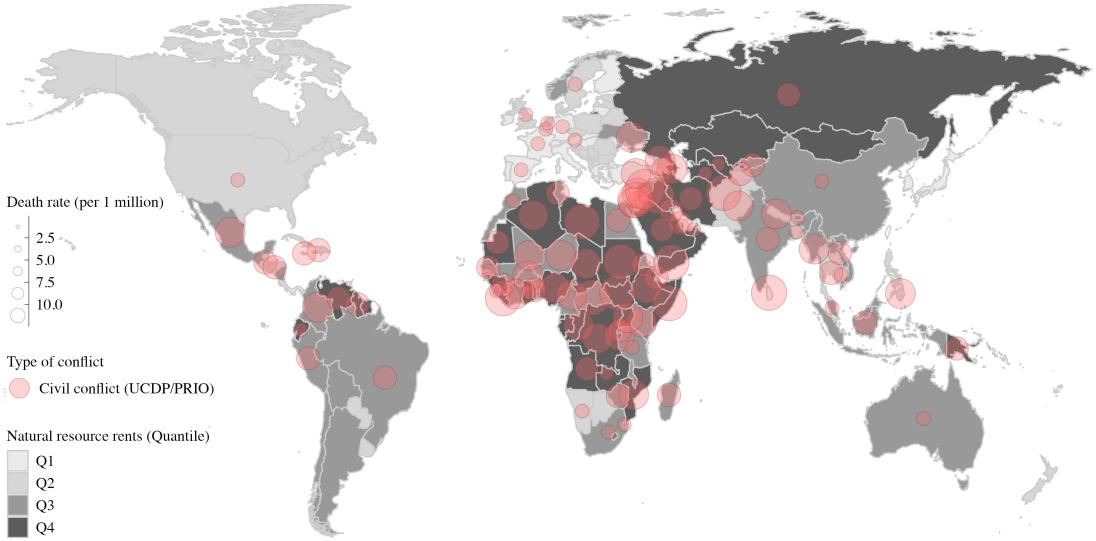
Table 5: The Effect of Police Violence on Electoral Performance of Incumbents

	Unconditional			Conditional		
	Elected	Vote Share	Runs	Elected	Vote Share	Margin
	(1)	(2)	(3)	(4)	(5)	(6)
Incumbency	-25.786***	-6.305***	-17.245***	-31.142***	-0.067	-6.850***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.838)	(0.003)
Incumbency × Police Violence	22.603	1.228	21.507	22.440	0.380	-1.496
	(0.194)	(0.893)	(0.426)	(0.196)	(0.888)	(0.841)
	[0.089]	[0.446]	[0.208]	[0.090]	[0.443]	[0.581]
Bandwidth	8.9	8.6	9.4	8.1	10.8	9.8
Dep. variable mean (%)	22.93	16.20	62.81	36.51	28.38	-4.42
Observations	2944	2944	2944	1849	1849	1849
Eff. Observations	1804	1760	1876	1067	1298	1224
Eff. Obs. (Police Violence = 1)	52	52	52	35	41	36

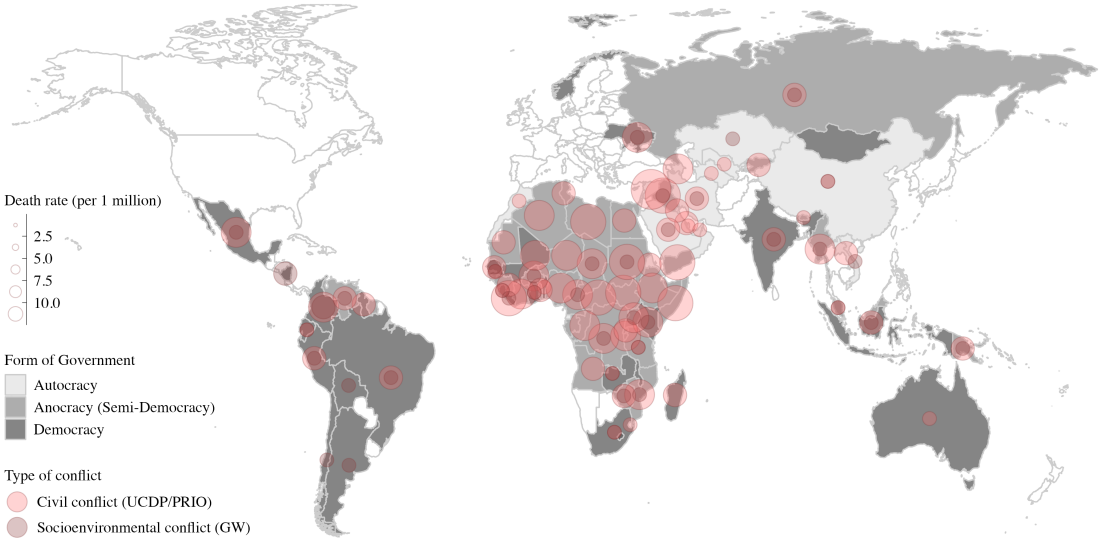
Notes: Local linear regression estimates with triangular kernel weights and optimal MSE bandwidth are reported. Robust p-values adjusted for clustering at the regional level are in parentheses. Robust *one-sided* p-values ($H_0 : \beta \leq 0$) adjusted for clustering at the state level are in brackets. The number of observations in the sample and the bandwidth, the optimal MSE bandwidth, and the dependent variable mean (in percent) are reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Figures

Figure 1: Natural Resource Rents and Socioenvironmental Conflict



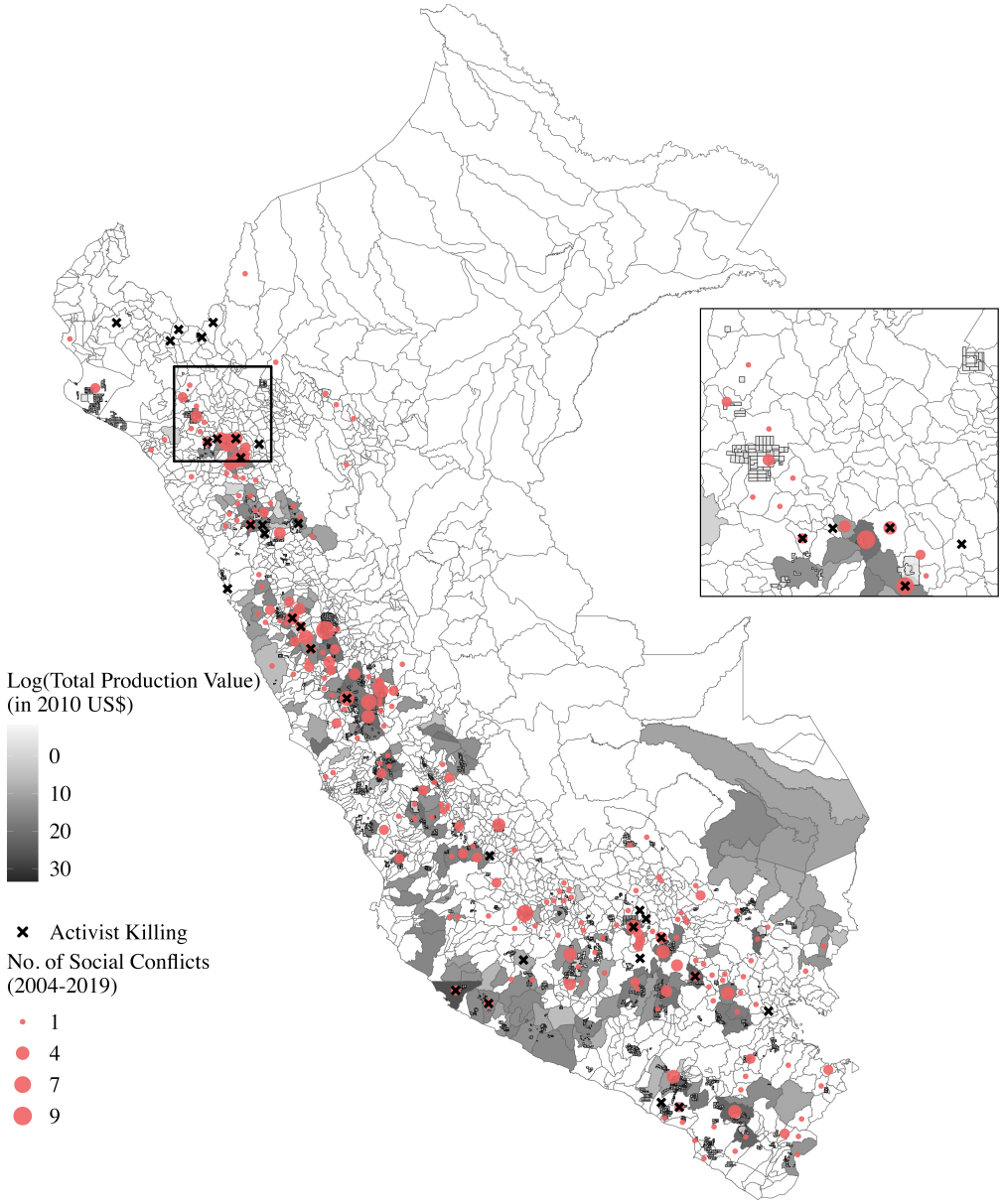
(a) Natural Resource Rents and Civil Conflict



(b) Socioenvironmental vs. Civil Conflict

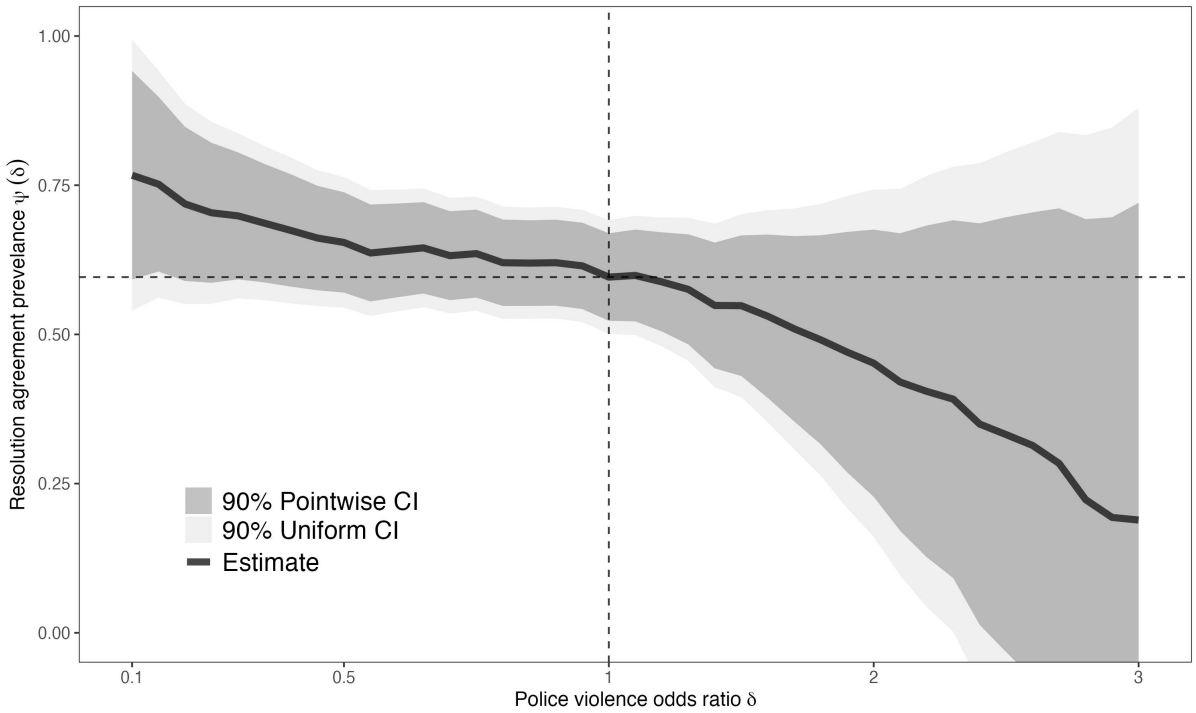
Notes: All figures are calculated for the period 2002-2019. Panel A presents the quantile distribution of the average mineral rents (as % of GDP) by country in sequentially darker shades of gray. Panel B is restricted to countries with above-median natural resource rents and displays their form of government based on their *Polity2* score ranging from -10 to $+10$ (Marshall et al., 2002): democracy ($> +5$), anocracy ($\leq +5$ and ≥ -5), and autocracy (< -5). "Best" fatality estimates in civil conflicts as registered in *UCDP GED* v22.1 (Sundberg and Melander, 2013; Davies et al., 2022) are presented as light-red shaded points. Panel B additionally displays in dark-red the total number of killed "land and environmental defenders" as reported by *Global Witness*. The scale of each point feature represents the inverse hyperbolic sine of the total number of deaths divided by the average population size in a country times one million. Data on population size and natural resource rents are taken from the *World Bank's WDI*.

Figure 2: Industrial Mining and Social Conflict



Notes: Boundary limits of municipalities—the third and lowest administrative level in Peru—are depicted. Mining concession areas are presented by the black-framed white areas.

Figure 3: Estimated Incremental Effect Curve for Resolution Probability



Notes: The estimated proportion of social conflicts ending in a resolution agreement $\psi(\delta)$ is shown when the odds of violent state suppression were to be multiplied by factor δ. $\delta = 1$ reports the resolution agreement prevelance for no intervention in the treatment process (but an intervention to remove right-censoring). *Pointwise* 90% confidence bands that allow valid inference at a single δ value are displayed in darker grey. *Uniform* 90% confidence bands to draw inference across the entire δ range are displayed in light grey.

Supplementary Online Appendix

Table of Contents

A Data Appendix **1**

A.1 Social Conflicts . . . 1

A.1.1 Social Conflict Data Set . . . 1

A.1.2 Protest Data Set . . . 2

A.1.3 Additional Figures . . . 3

A.2 Government Plans . . . 6

A.2.1 Classification of Government Plans . . . 6

A.3 Mineral Prices . . . 10

A.4 Variable Definitions . . . 12

B Mineral Rents and Social Conflict Violence **19**

B.1 Descriptive Statistics . . . 19

B.2 Robustness – Omitted Variables . . . 21

B.3 Robustness – Measurement . . . 23

B.3.1 Main Mineral Price vs. Price Index . . . 23

B.3.2 Outcomes . . . 24

B.3.3 Producing municipalities . . . 26

B.4 Robustness – Econometric Specification . . . 27

B.4.1 Levels vs. Differences . . . 27

B.5 Robustness – Additional Sensitivity Checks . . . 32

B.5.1 Seasonality . . . 32

B.5.2 World Market Share . . . 33

B.5.3 Spatial Kernel . . . 34

B.5.4 Multiple Hypothesis Correction . . . 35

C Pro-Mining Local Politicians, State Repression, and Rent Seeking **37**

C.1 Manipulation Test . . . 37

C.2 Additonal RDD Results . . . 38

D Democratic Accountability **42**

D.1 Additional Results 42

E Is Violent State Repression Effective? 43

E.1 Descriptive Statistics 43

E.2 State Repression and Social Conflict Duration 44

E.3 Dynamic Causal Inference and Incremental Causal Effects 47

E.3.1 Technical Appendix 47

E.3.2 Additional Results 51

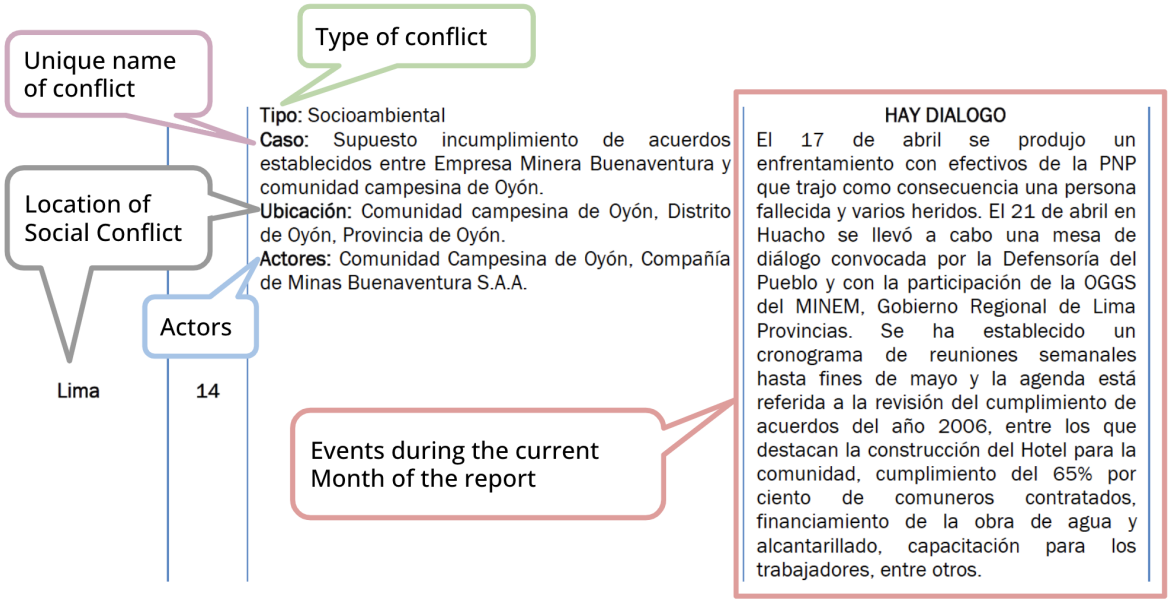
A Data Appendix

A.1 Social Conflicts

A.1.1 Social Conflict Data Set

Peruvian Ombudsman Since April 2004, the Office of the Ombudsman (*Defensoría del Pueblo de Perú*) has published a monthly report on social conflicts in Peru. Each report follows a fairly consistent format.[41]

Figure A.1: Ombudsman Report Example Entry



Source: Defensoría del Pueblo de Perú "Reporte de Conflictos Sociales Nº 86" (April 2014).

Mining Project Ownership I rely on Bureau van Dijk's *Orbis* database to obtain corporate ownership information. *Orbis* reports shareholder history, allowing me to trace ownership of mining subsidiaries over the sample period from March 2004 to December 2019. I cross-validate this information against publicly available reports of corporations and authorities (e.g., SEC), the annual *USGS Mineral Yearbook* "Mineral Industry of Peru" reports and S&P's SNL Metals and Mining dataset. The latter, moreover, provides information on the ownership of mining projects and concessions as well as their locations, allowing me to trace ownership even when no company name is noted in the reports.

[41] The ombudsman has updated the report's outline over time, although key formats and information were preserved over the sample period. The report numbers for which the ombudsman made nonnegligible updates to the report outline are as follows: 3, 9, 11, 21, 35, 50, 72, and 90.

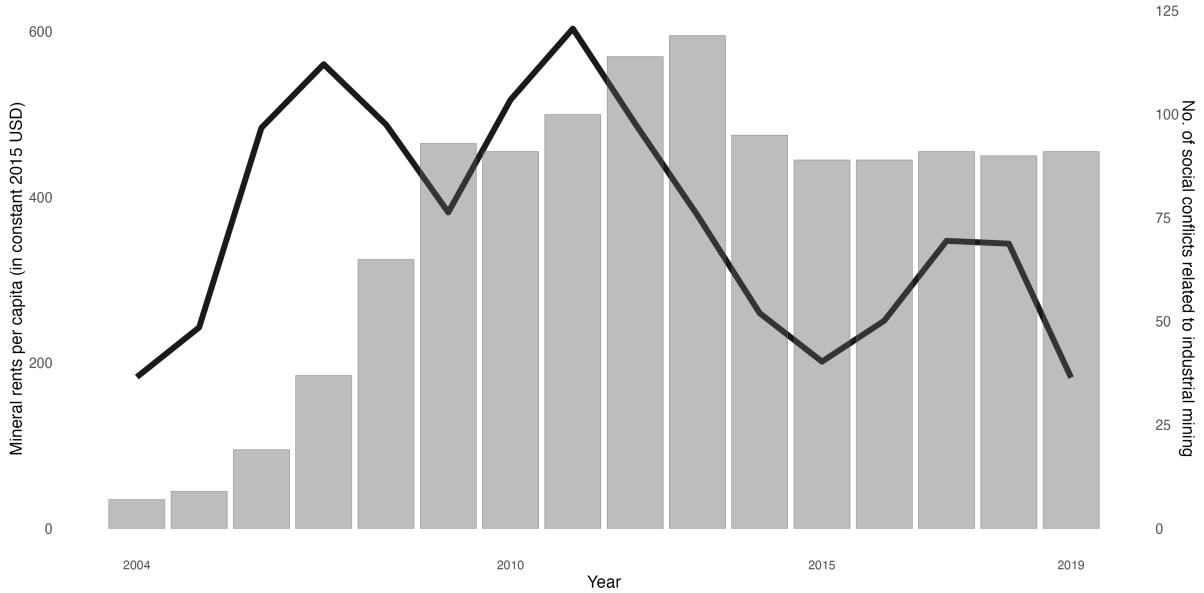
Geoprecision Code The highest precision level, 1, is recorded as the default value of local social conflicts at the municipality level (ADMIN3). If the social conflict is noted to take place at the provincial (ADMIN2) or departmental (ADMIN1) level, the precision level is 2 or 3, respectively, and the province capital or department capital is used, unless previous or subsequent reports have information at a more granular level. In the latter case, the noted municipality- and/or province-level information in other reports is used.

A.1.2 Protest Data Set

Geoprecision Code If the report notes a particular town or municipality (ADMIN3), the highest precision level, 1, is recorded. If the source material notes the area around the mine or community, the geoprecision code is 2. The same code is applied if the mentioned location of the protest covers an area comprising multiple municipalities. If only the province (ADMIN2) is mentioned, the provincial capital is used, and the precision level is 3. If only the department and no other information is available, the departmental capital is used and noted with precision level 4.

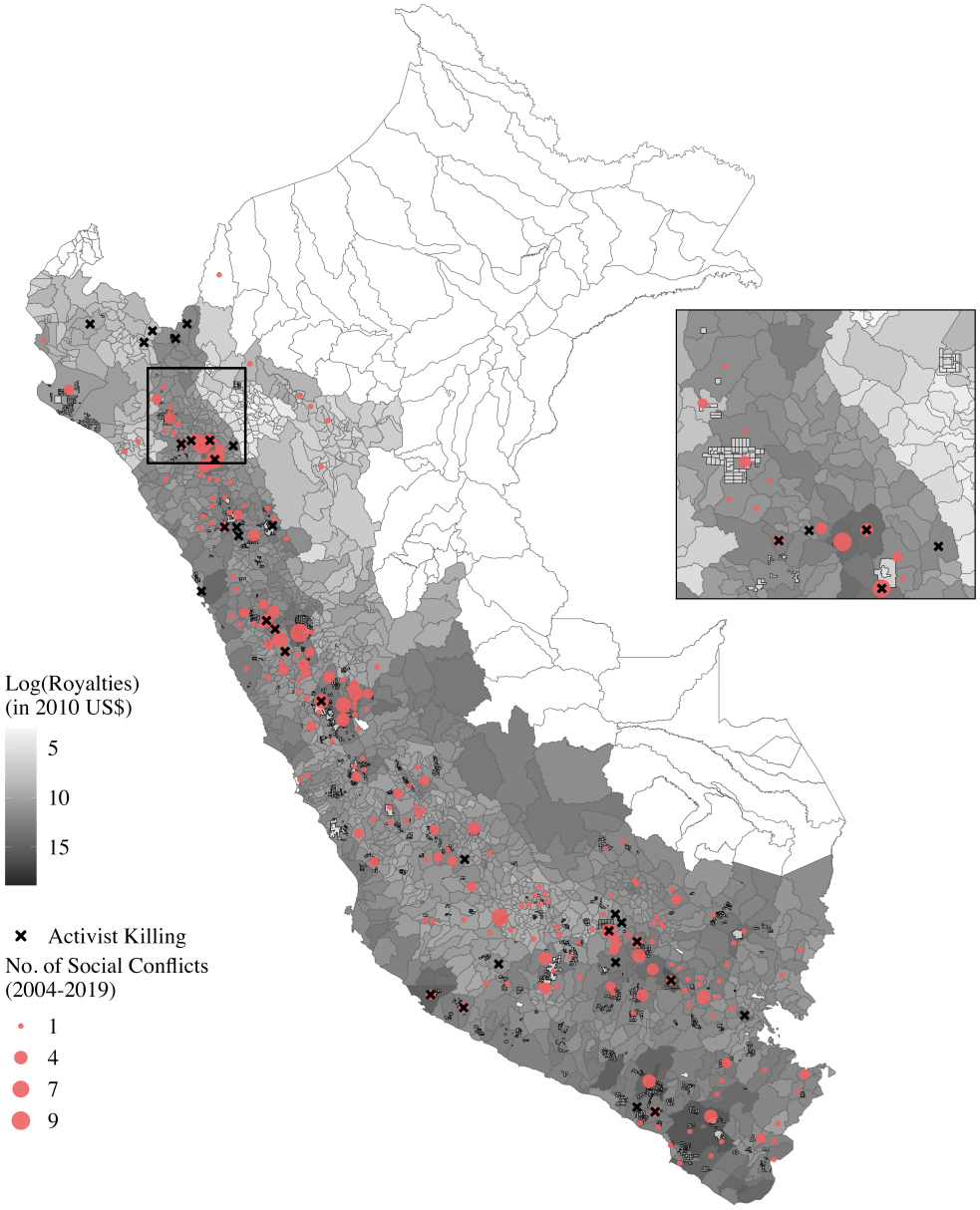
A.1.3 Additional Figures

Figure A.2: Mineral Rents and Social Conflict (2004–2019)



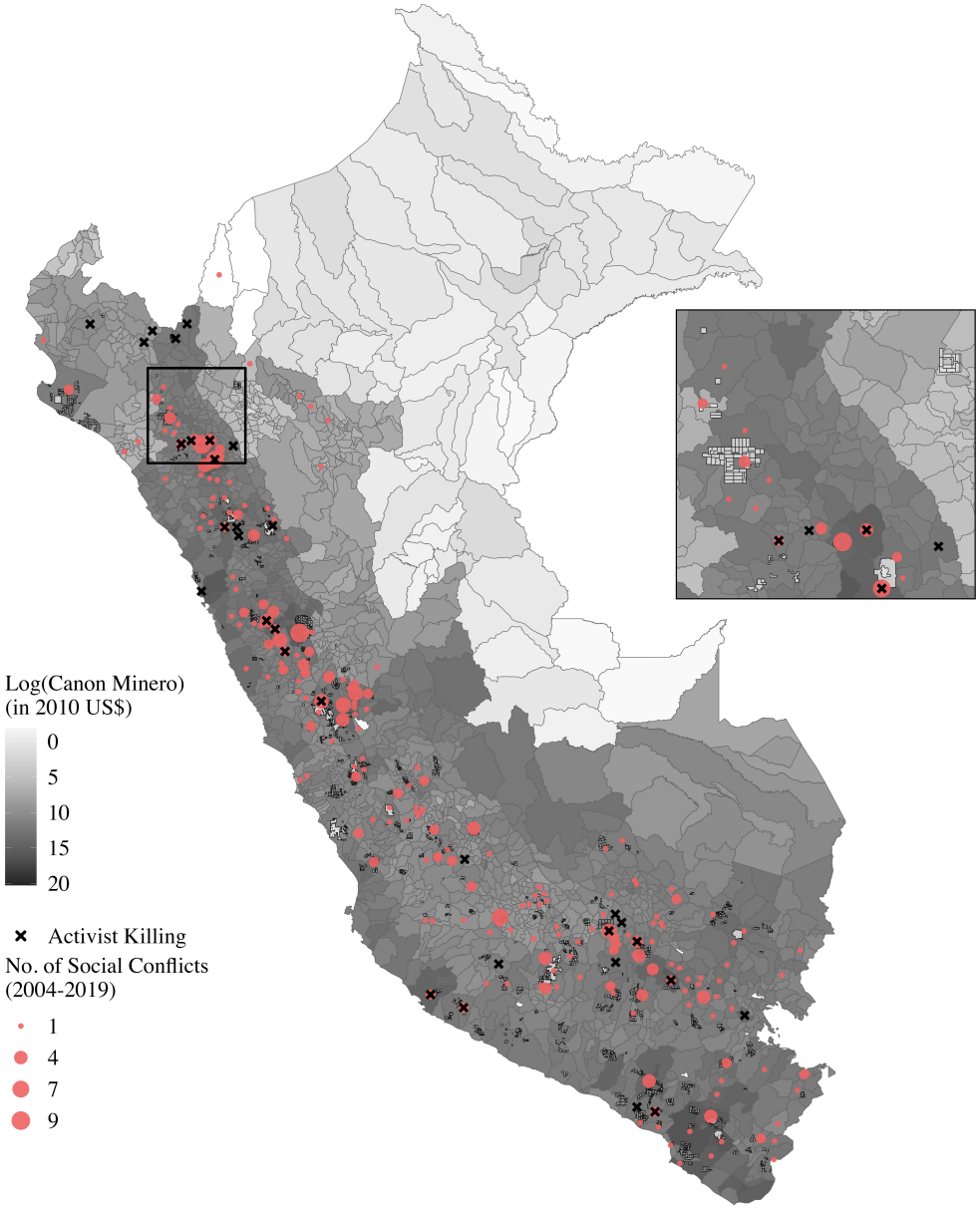
Notes: Mineral rents per capita (in constant 2015 USD) are computed from the *World Bank's World Development Indicators* and depicted by the black solid line. Grey bars display the number of social conflicts related to industrial mining and are calculated by the author on the basis of "*Conflictos sociales*" reports published by the Peruvian Office of the ombudsman (*Defensoria del Pueblo*).

Figure A.3: Social Conflict and Royalties



Notes: Boundary limits of municipalities–the third and lowest administrative level in Peru–are depicted. Mining concession areas are presented by the black framed white areas. Total transfers received by each municipality from mining royalties in real 2010 US dollars over the period from 2002-2019 are depicted.

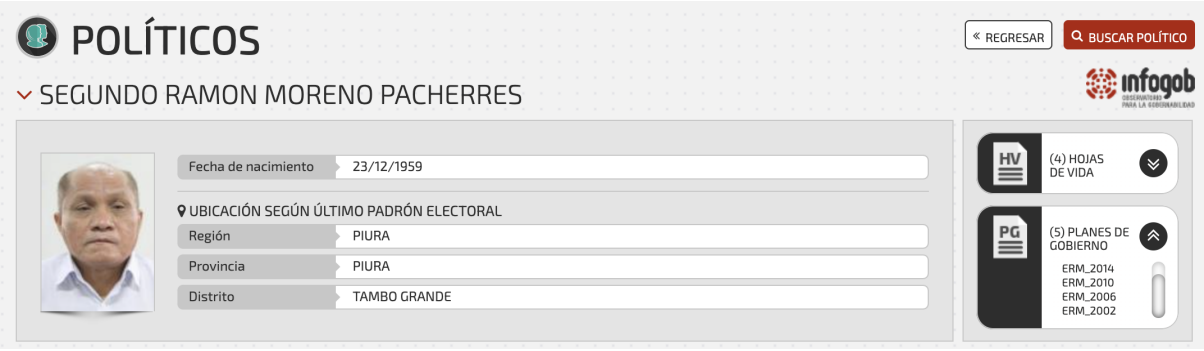
Figure A.4: Social Conflict and Canon Minero



Notes: Boundary limits of municipalities–the third and lowest administrative level in Peru–are depicted. Mining concession areas are presented by the black framed white areas. Total transfers received by each municipality from *canon minero* in real 2010 US dollars over the period from 2002-2019 are depicted.

A.2 Government Plans

Figure A.5: Mayoral Candidate Government Plans



(a) Infogob profile of mayoral candidate

De esta manera el nuevo Municipio con sus autoridades ediles del período 2003 - 2006 , construirán un Municipio honesto, leal y capaz, poniendo a la persona como fin supremo de la sociedad y del Estado, un Municipio eficiente - productivo con una real vocación de servicio, que apueste por fortalecer la democracia y la descentralización en nuestra Patria, a través de la participación ciudadana y comunal..

Por esto y mucho más, presentamos a nuestro noble pueblo de Tambogrande y caseríos, nuestro Plan de Gobierno Municipal 2003 - 2006, bajo su lema "Desarrollo del Agro sin Minería".

(b) Government plan excerpt

Notes: The top panel presents the profile of a mayoral candidate in the *Infogob* database. The bottom panel displays an excerpt of the mayoral candidate's government plan for 2002. All available government plans can be accessed in the *planes de gobierno* window displayed at the bottom right corner in the top panel.
Source: infogob.jne.gob.pe/Politico/FichaPolitico/segundo-ramon-moreno-pacherres_historial-partidario_LqJgXVKnPC4=JV (last accessed 5 September 2023).

A.2.1 Classification of Government Plans

The classification procedure can be broadly summarized into these three (four) steps:

1. Preprocess government plans to obtain text corpus.
2. Filter government plans that (i) deal with mining and (ii) have relevant passages concerned with mining.
3. Ask GPT-4 to answer 5 questions based on the passages extracted in step 2 to determine the candidate's sentiment toward formal mining.

- If the answers are conflicting, feed text back to GPT-4 again with an adjusted prompt tailored toward resolving the conflict.

(4) If the classification procedure returns an error in any of the first three stages of the process, label the government reports manually.

Below, I elaborate on each of these steps.

Preprocessing I use the `python` libraries `pdfplumber` and `pytesseract` to extract text from pdf documents and scans and the `Linux` application `antiword` to process `.doc` Word documents. The text output is saved in a simple `.txt` file.

Filtering I use a semiautomatically generated keyword list to identify relevant mining terms in a text document. In particular, I use the pretrained Spanish word embeddings from Cardellino (2019) to obtain the top 50 most similar words (in cosine similiarity) to "*minera*". I clean the top 50 results of ambiguous (e.g., *minerales*) or irrelevant (e.g., *petrolera*) words and supplement the list with terms to identify mining corporations active in Peru (e.g., *bhp*). The final keyword list is: *carbonífera*, *barrick*, *yacimiento*, *copper*, *riotinto*, *bauxita*, *siderúrgico*, *minería*, *minero*, *mineros*, *aurífera*, *codelco*, *carbón*, *mineras*, *metalúrgicas*, *boliden*, *mina*, *glencore*, *siderúrgicas*, *siderurgia*, *mines*, *minas*, *mining*, *minera*, *spcc*, *plata*, *copper*, *zinc*, *minsur*, *shougang*, *silver*, *hudbay*, *marcobre*, *nexa*, *gold*, *shahuindo*, *coal*, *antamina*, *xstrata*, *newmont*, *bhp*, *doe*.
With the keyword list in hand, I use `spaCy` (Honnibal et al., 2020) to segment the text of each government plan document into sentences. I subsequently split each sentence into separate tokens (words) and compare each token against the keyword list.[42] I retain all sentences that include at least one keyword and a window of the 4 sentences before and after the sentence containing the keyword to retain contextual information. Finally, I combine all selected sentences to obtain a concise representation of the *relevant* passages of the original document.[43] Documents that do not contain any mining keywords are labeled "unknown."

GPT-4 The relevant passages for each government plan obtained in the previous step (denoted `.text_esp` below) are then used to send the following prompt to GPT-4:

```
[.text_esp]

From the text in Spanish above, answer the following questions.
Report results in a json array with a json object with three keys for each
    question: answer in English; reasoning in English; supporting quotes.
```

[42] I apply some simple preprocessing to trim excess whitespace and remove punctuation in each sentence.

[43] Note that I remove duplicates if the window of one sentence contains part of the window of another keyword sentence.

```
Provide reasoning up to 50 words.

* Does the text focus on the positive aspects of formal mining? Answer yes/
    no.
* Does the text suggest the promotion of formal mining activity? Answer yes
    /no.
* Does the text focus on the negative impacts of formal mining? Answer yes/
    no.
* Does the text suggest restrictions on formal mining activity? Answer yes/
    no.
* Is the text neutral towards formal mining? Answer yes/no.
```

The response from GPT-4 takes the following standardized form, as illustrated in the example output below:

```
[{
    "answer": "no",
    "reasoning": "The text does not focus on the positive aspects of formal
        mining.",
    "supporting_quotes": ""
},
{
    "answer": "no",
    "reasoning": "The text does not suggest the promotion of formal mining
        activity.",
    "supporting_quotes": ""
},
{
    "answer": "yes",
    "reasoning": "The text mentions the need for environmental impact
        studies and damage mitigation due to mining exploitation.",
    "supporting_quotes": "ejecutar estudios de impacto ambiental y
        mitigaci n de da os por la explotaci n de los recursos mineros de
        cuajone y quellaveco pr ximo a explotar."
},
{
    "answer": "yes",
    "reasoning": "The text suggests the implementation of an environmental
        management plan and the declaration of intangibility of certain
        areas.",
    "supporting_quotes": "plan de gesti n ambiental. declaraci n de
        intangibilidad de los arcos glaciares de arondaya."
},
{
    "answer": "no",
    "reasoning": "The text is not neutral towards formal mining as it
        highlights the need for environmental impact studies and damage
        mitigation.",
```

```
    "supporting_quotes": "ejecutar estudios de impacto ambiental y
        mitigaci n de da os por la explotaci n de los recursos mineros de
         cuajone y quellaveco pr ximo a explotar."
}]
```

The five questions are designed to allow classification of the government plan's sentiment toward mining—specifically, as *pro*, *anti*, or *neutral*. For instance, I classify a mayoral candidate as *pro*-mining if at least one of the first two questions is answered with "yes" (`Y`) by GPT-4 and the answers to questions three, four, and five are "no" (`N`). Below, I outline in detail how the answers translate to labels:

1. *pro*: `YYNNN`, `YYNYN`, `YNNNN`, `NYNNN`.
2. *anti*: `NNYYN`, `NYYYN`, `NNYNN`,`NNNYN`.
3. *neutral*: `NNNNY`, `NNNNN`.
4. *conflict*: otherwise

Government plans for which a logical conflict in the answers exists—e.g., if the focus is on both the benefits and negative impacts of formal mining (`YNYNN`)—the plans are fed back to GPT-4 with the following adjusted prompt:

```
[.text_esp]

From the text in Spanish above, answer the following question. Report
    results in a json array with a json object with three keys for each
    question: answer in English; reasoning in English; supporting quotes.
    Provide reasoning up to 50 words.
* Does the text focus on the benefits and promotion of formal mining or on
    the negative impact and restriction of formal mining? Answer benefits
    and promotion/negative impact/neither.
```

Errors If the classification procedure returns an error in any of the first three stages of the process, e.g., because the pdf document is protected, the government report is manually inspected and labeled by the principal investigator. Documents that cannot be classified due to the document being password protected or containing only a blank page are treated as if missing.

Imputation If a candidate runs in *multiple* mayoral races but the government plan for a particular year is missing due to either the document being password protected, a blank page, or not available in the *Infogob* database, I impute the sentiment from the remainder of available plans. Note that I require that the sentiment of the available government plans for a candidate is consistent, i.e. does not change over time and remains either anti, pro, or neutral. If the sentiment changes over time, the government plan continues to be treated as missing.

A.3 Mineral Prices

Mineral price information at monthly (and yearly) frequency is retrieved from the *World Bank Commodity Price Data*.[44] The price development of each mineral over the period 2002–2020 is graphed in Figure A.6. At yearly frequency, I supplement the data set with digitized price information on minerals mined in Peru but not covered by the *World Bank* from the *USGS*'s "Mineral Commodity Summaries" or "Mineral Yearbooks."[45] Note that the *USGS* mineral prices do not reflect the world mineral price but the average price for the US. All mineral prices are converted to real USD with the MUV Index provided by the *World Bank* and uniformly expressed as $\$/kg$. A list of the minerals mined in Peru and priced at monthly and yearly frequency is provided below:

Monthly Mineral Prices:

- *World Bank*: Iron, Copper, Lead, Tin, Zinc, Gold, Silver.

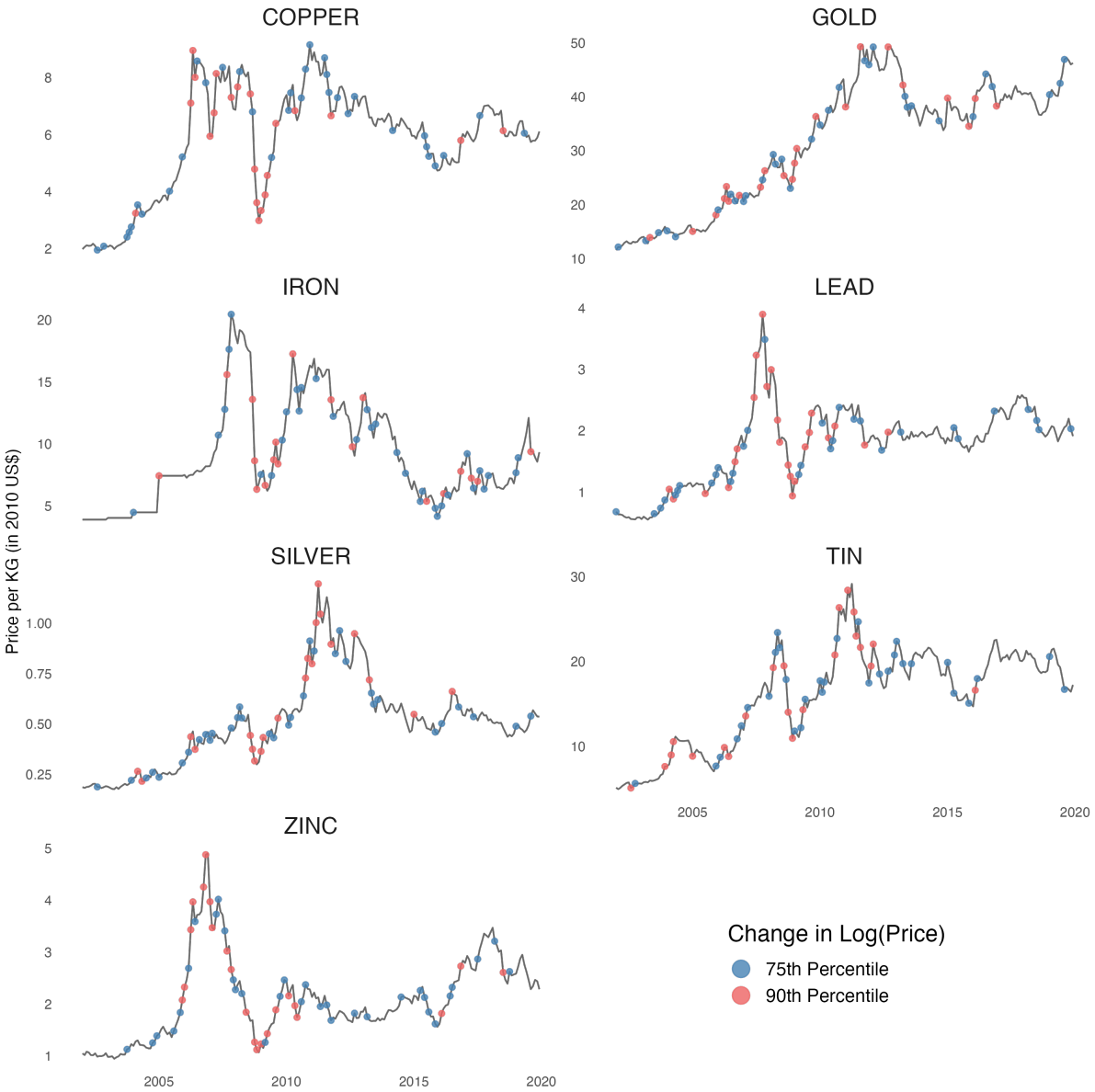
Yearly Mineral Prices:

- *World Bank*: Iron, Copper, Lead, Tin, Zinc, Gold, Silver.
- *USGS*: Arsenic, Bismuth, Cadmium, Manganese, Molybdenum, Tungsten.

[44] For more details, see worldbank.org/en/research/commodity-markets#1.

[45] For more details, see usgs.gov/centers/national-minerals-information-center/commodity-statistics-and-information.

Figure A.6: Mineral Prices



Notes: Prices for each mineral are expressed in real 2010 USD per kg for the period from January 2002 to December 2020.

A.4 Variable Definitions

Table A.1: Variable Definitions

Variable	Description
Section 2: Mineral Rents and Social Conflict Violence	
Arrests	Dummy equaling 1 if at least one arrest of a protester was observed in municipality i in month t and 0 otherwise.
Injuries	Dummy equaling 1 if at least one injury of a protester was observed in municipality i in month t and 0 otherwise.
Casualties	Dummy equaling 1 if at least one casualty among protesters was observed in municipality i in month t and 0 otherwise.
Killing	Dummy equaling 1 if at least one casualty among activists was observed in municipality i in month t and 0 otherwise. This can include fatal violence during protests or assassinations.
Protest	Dummy equaling 1 if at least one protest was observed in municipality i in month t and 0 otherwise.
Protester riots	Dummy equaling 1 if at least one incident of violence or destruction of property by protesters was observed in municipality i in month t and 0 otherwise.
Protester violence	Dummy equaling 1 if at least one incident of protester violence was observed in municipality i in month t and 0 otherwise.
Main mineral price	Main mineral price in month t for municipality i. The main mineral in a municipality is determined by *(i)* total production value or *(ii)* the primary commodity count of concessions (in nonproducing municipalities).
Price index	Weighted price index in month t for municipality i. Weights are computed as each mineral's share of *(i)* total production value or *(ii)* the primary commodity count of concessions (in nonproducing municipalities).
Main crop price	Crop price in month t of main crop in municipalities i. The main crop in municipality i is determined by the total production value of each crop with monthly price data coverage from the *World Bank*. Crop production for the year 2000 is obtained from *EarthStat* database (Monfreda et al., 2008).

Table A.1: Variable Definitions *(continued)*

Variable	Description
Temperature	Average temperature in month t in municipality i. Data on temperature are obtained from the *CRU TS v4* database (Harris et al., 2020).
Rainfall	Total precipitation in month t in municipality i. Data on monthly precipitation are obtained from the NOAA Global Precipitation Climatology Centre.
Section 3.1: Pro-Mining Local Politicians, State Repression, and Rent Seeking	
Change in police violence over term	Difference between post-election average of binary term-year indicators equaling one if at least one incident of police violence against protesters was registered in the municipality and a binary indicator of police violence in the year prior to the mayor's election.
Change in police violence frequency over term	Difference between post-election average of the share of months with police violence in each term-year and share of months with police violence in the year prior to the mayor's election.
Change in experienced corruption	Difference between post-election average in the percentage of respondents in each-term year that stated that at least one member of the surveyed household "was requested to, felt obligated to or voluntarily gave a bribe to a public official in the last 12 months" relative to the total number of respondents in the municipality and the percentage in the year prior to the mayor's election. Data on responses is obtained from the *Encuesta Nacional de Hogares (ENAHO)* survey on local governance.
Probability of corruption investigation against mayor	Dummy equal to 1 if at least one investigation for crimes in office was launched against the mayor during or after their term. Data on corruption investigations and charges are drawn from Pacheco Palacios et al. (2022).
Probability of *initial investigation*	Dummy equal to 1 if at least one initial investigation for crimes in office was launched against the mayor but none of the investigations has progressed to the formal legal stage (stage 2 to 3). Data on corruption investigations and charges are drawn from Pacheco Palacios et al. (2022).

Table A.1: Variable Definitions *(continued)*

Variable	Description
Probability of *legal persecution*	Dummy equal to 1 if at least one investigation for crimes in office against the mayor has progressed to the formal legal stage (stage 2 to 3). Data on corruption investigations and charges are drawn from Pacheco Palacios et al. (2022).
Win margin	Margin in votes between pro-mining and anti-mining candidate; > 0 if the pro-mining candidate wins the election, and < 0 if the anti-mining candidate wins the election. Data on election outcomes are obtained from the *National Jury of Elections (JNE).*
Political competition	Political competition is calculated as the inverse of the sum of squared vote shares of each candidate within an electoral race (Artiles et al., 2021). Data on election outcomes are obtained from the National Jury of Elections (JNE).
Turnout	Share of eligible voters that voted in municipal election. Data on election outcomes are obtained from the National Jury of Elections (JNE).
Canon minero	*Canon minero* transfer to municipality in year prior to election.
Royalties	Royalty transfer to municipality in year prior to election.
Area	Surface area of municipality (in km^2).
Nighttime lights	Average nighttime lights in municipality in year prior to election. For consistency over time, the harmonized global nighttime light dataset of Li et al. (2020) is used.
River length	Total length of rivers in municipality (in km). Data on waterways are obtained from the *HydroSheds* database.
Elevation	Mean elevation of municipality. Elevation data are obtained from *SRTM* v4 dataset.
Lake area share	Ratio of land covered by lakes to total area in municipality. Data on lake size are obtained from the *HydroSheds* database.
Indigenous	Dummy equal to 1 if the municipality has at least one indigenous community or native land. Data on the number of indigenous citizens and native lands is taken from the *Mapeo territorial* of the *Instituto del Bien Común.*

Table A.1: Variable Definitions *(continued)*

Variable	Description
Population density	Municipal population per km^2 (population count divided by municipality area) in the year prior to the election. Data on population counts are retrieved from the official Peruvian census in 1993, 2007, and 2017; Population counts are linearly interpolated between census waves (extrapolated after 2017).
Road density	Road length (in km) divided by municipality area (km^2). Data on the Peruvian road network are obtained from the *Ministerio de Transportes y Comunicaciones*.
Provincial capital	Dummy equal to 1 if the municipality contains the provincial capital.
ENAHO Respondents	Total number of respondents situated in the municipality that answered the "" question in the ENAHO survey.
Section 3.2: Democratic Accountability	
Incumbency	Dummy equal to 1 if the candidate is the incumbent mayor.
Police Violence	Dummy equal to 1 if at least one incident of police violence in the municipality was registered during the mayors term.
Unconditional win probability	Dummy equal to 1 if the candidate (incumbent or runner-up) wins in the next election. Candidates who do not not run or run and loose are coded as 0.
Unconditional vote share	Vote percentage of candidate (incumbent or runner-up) in the next election. The vote share of candidates who do not run in the next election is coded as 0.
Probability of running in next election	Dummy equal to 1 if the candidate (incumbent or runner-up) runs in the next election.
Conditional win probability	Dummy equal to 1 if the candidate (incumbent or runner-up) ran in the next election and won. Candidates who do not run in the next election are dropped from the sample.
Conditional vote share	Vote percentage of candidate (incumbent or runner-up) in the next election if candidate ran again. Candidates who do not run in the next election are dropped from the sample.

Table A.1: Variable Definitions *(continued)*

Variable	Description
Conditional victory margin	If the candidate (incumbent or runner-up) does not win in the next election, the victory margin is defined as the candidate's own vote share minus the vote share of the election winner. Candidates who do not run in the next election are dropped from the sample.
Section 3.3: Is Violent State Repression Effective?	
Resolution	Dummy equaling 1 if the conflict ended with an official resolution agreement between conflict parties.
Removed	Dummy equaling 1 if the conflict ended with a removal from the watch list of the Peruvian ombudsman due to prolonged inactivity.
Right-censored	Dummy equaling 1 if the end of the conflict was not observed either *(i)* due to the *unexplained* removal of the conflict from the watch list of the Peruvian ombudsman or *(ii)* due to the end of the sample period in Dec 2019.
Police violence	Binary *treatment* indicator equaling 1 if at least one incident of police violence against protesters in connection with social conflict i was observed in conflict-year t.
No. of months w/ protests	Number of months with at least one protest in connection with social conflict i in conflict-year t.
No. of months w/ protester riots	Number of months with at least one riot of protesters in connection with social conflict i in conflict-year t.
Δ Main mineral price	Change over conflict year t in the main mineral prices averaged across municipalities where social conflict i is located. The main mineral in a municipality is determined by the *(i)* total production value or *(ii)* primary commodity count of concessions (in nonproducing municipalities).
Canon minero	Sum of *canon minero* transfers in conflict-year t received by all municipalities where social conflict i is located.
Royalties	Sum of *royalty* transfers in conflict-year t received by all municipalities where social conflict i is located.
Nighttime Lights	Average of (mean) nighttime lights conflict-year t across municipalities where social conflict i is located. For consistency over time, the harmonized global nighttime light dataset from Li et al. (2020) is used.

Table A.1: Variable Definitions *(continued)*

Variable	Description
Duration	Duration of social conflict in months (Section E.2) or conflict-years (Section 3.3.2).
Indigenous	Dummy equal to 1 if the municipalities where social conflict i is located have any indigenous communities or native lands. Data on the number of indigenous citizens and native lands is taken from the *Mapeo territorial* of the *Instituto del Bien Común*.
Population density	Total population at the start of social conflict i in municipalities where social conflict i is located divided by the total surface area (in km^2) of these municipalities. Data on population counts are retrieved from the official Peruvian census in 1993, 2007, and 2017; population counts are linearly interpolated between census waves (extrapolated after 2017).
Elevation	Average (mean) elevation across municipalities where social conflict c is located. Elevation data are obtained from the *SRTM* v4 dataset.
Road density	Total road length (in km) in municipalities where social conflict i is located divided by total surface area (in km^2) in these municipalities. Data on the Peruvian road network are obtained from the *Ministerio de Transportes y Comunicaciones*.
Share of area covered by lakes	Ratio of land covered by lakes to total area in municipalities where social conflict i is located. Data on lake sizes are obtained from the *HydroSheds* database.
River density	Total length of rivers (in km) where social conflict i is located divided by total surface (in km^2) are in these municipalities. Data on waterways are obtained from the *HydroSheds* database.
Area	Total surface area (in km^2) in municipalities where social conflict i is located.

Table A.1: Variable Definitions *(continued)*

Variable	Description
Foreign ownership	Dummy equal to 1 if (at least one of) the majority owner(s) of the mine/project associated with social conflict i is (are) headquartered outside of Peru. Historical ownership shares are obtained from Bureau van Dijk's *Orbis* database and cross-validated with annual reports if available. Location information is obtained from *Orbis* and S&P's *Compustat* database.
Market value of majority owner	Market value (in US$) of the majority owner(s) of the mine/project associated with social conflict i (measured at the start of social conflict i). The average market value is used if there is more than one publicly traded majority owner. Data on market capitalization are obtained from S&P's *Compustat* database.

B Mineral Rents and Social Conflict Violence

B.1 Descriptive Statistics

Table B.1: District Government Revenues from Royalties and *Canon Minero* (2002–2019)

		N	Mean	Median	SD	Min	Max
All	*Canon Minero* (in 2010 USD)	1873	6,213,556	941,625	22,433,363	0	535,857,094
	Royalties (in 2010 USD)	1873	1,278,833	236,485	5,005,074	0	77,952,062
Production	*Canon Minero* (in 2010 USD)	215	18,701,791	4,811,426	48,552,674	4,275	535,857,094
	Royalties (in 2010 USD)	215	4,580,338	1,153,231	11,737,776	0	77,952,062
Concessions	*Canon Minero* (in 2010 USD)	281	10,140,780	2,077,637	28,146,551	599	314,122,841
	Royalties (in 2010 USD)	281	1,912,803	451,413	5,707,105	472	72,085,687
None	*Canon Minero* (in 2010 USD)	1377	3,462,270	524,076	11,033,366	0	198,282,741
	Royalties (in 2010 USD)	1377	633,976	150,223	1,955,571	0	39,808,289

Notes: Author's computation on the basis of data from the *Ministerio de Economía y Finanzas* (MNF).

Table B.2: World Production Shares by Commodity

	(1) Copper	(2) Gold	(3) Iron	(4) Lead	(5) Silver	(6) Tin	(7) Zinc
Panel A: Peru (Total)							
2002	6.2	5.4	0.3	10.0	13.4	**18.5**	13.2
2003	6.1	6.6	0.3	10.4	14.8	**18.4**	13.9
2004	7.1	7.1	0.3	9.7	15.5	**15.9**	12.5
2005	6.7	8.4	0.3	9.8	**16.5**	14.5	12.2
2006	6.9	8.3	0.3	9.0	**17.2**	12.6	12.0
2007	7.7	7.1	0.2	8.7	**16.8**	12.2	13.2
2008	8.2	8.0	0.2	9.0	**17.3**	13.0	13.8
2009	8.0	7.4	0.2	7.8	**17.7**	14.4	13.5
2010	7.9	6.4	0.2	6.3	**15.8**	12.8	12.2
2011	7.7	6.2	0.2	4.9	**14.6**	11.8	9.8
2012	7.7	6.0	0.2	4.8	**13.6**	10.9	9.5
2013	7.5	5.4	0.2	4.8	**14.1**	8.1	10.1
2014	7.5	4.7	0.2	5.7	**14.1**	8.1	9.9
2015	8.9	4.7	0.3	6.4	4.7	6.7	**11.1**
2016	11.7	4.9	0.6	6.7	**17.0**	6.5	10.6
2017	12.2	4.7	0.6	6.7	**16.0**	5.7	11.8
2018	12.0	4.3	0.6	6.3	**15.5**	5.8	11.8
2019	12.1	3.9	0.7	6.5	**14.6**	6.7	11.0
Panel B: Municipality (Max)							
2002	2.5	1.7	0.3	3.9	2.0	**18.5**	3.3
2003	1.9	2.0	0.3	3.7	1.9	**18.4**	4.0
2004	2.6	1.9	0.3	3.2	1.9	**15.9**	3.0
2005	2.6	3.4	0.3	3.3	2.0	**14.5**	2.8
2006	2.6	3.3	0.3	1.4	1.6	**12.6**	2.0
2007	2.2	2.0	0.2	1.7	1.7	**12.2**	3.1
2008	2.3	2.5	0.2	1.7	2.0	**13.0**	3.5
2009	2.2	2.6	0.2	1.1	2.3	**14.4**	4.6
2010	2.1	1.8	0.2	0.7	2.1	**12.8**	3.6
2011	2.2	1.4	0.2	0.6	1.6	**11.8**	2.1
2012	2.7	1.2	0.2	0.6	1.6	**10.9**	2.0
2013	2.5	0.7	0.2	0.7	2.0	**8.1**	2.4
2014	2.0	0.7	0.2	0.7	1.6	**8.1**	2.0
2015	2.2	0.9	0.3	0.7	0.7	**6.7**	2.3
2016	2.6	0.6	0.6	0.7	2.5	**6.5**	2.1
2017	2.5	0.5	0.6	0.7	2.3	**5.7**	3.5
2018	2.4	0.5	0.6	0.6	2.0	**5.8**	3.8
2019	2.3	0.5	0.7	0.6	1.8	**6.7**	2.8

Notes: Data on world production of minerals is obtained from the USGS *Mineral Commodity Summaries*. Production data by mineral for Peru at the national and municipality levels are obtained from USGS and MINEM. The largest share in a year is highlighted in bold.

B.2 Robustness – Omitted Variables

Table B.3: Neighborhood Analysis – 10 Nearest Neighbors

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
ln(Price)	0.0000	-0.0005	-0.0002**	-0.0003	-0.0004
	(0.0001)	(0.0003)	(0.0001)	(0.0002)	(0.0002)
$M\times$ ln(Price)	0.0011**	0.0022**	0.0007*	0.0005	0.0005
	(0.0005)	(0.0009)	(0.0004)	(0.0004)	(0.0005)
Neighbor × year FEs	✓	✓	✓	✓	✓
Month FEs	✓	✓	✓	✓	✓
Observations	564110	564110	564110	564110	564110

Notes: M equals one for mining districts (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a district is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Robust standard errors are clustered at the neighbor and month level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Additional Time-Varying Controls

	Force used against protesters												Protester behavior							
	Arrests				Injuries				Casualties				Violence				Riots			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
$M\times$ ln(Price)	0.0004*	0.0004**	0.0003	0.0003	0.0019**	0.0019***	0.0018**	0.0018**	0.0006**	0.0007**	0.0006**	0.0006**	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Ln(Crop price)	-0.0001			0.0000	0.0001			0.0002	0.0003			0.0004	-0.0001			0.0000	-0.0001			0.0000
	(0.0004)			(0.0004)	(0.0009)			(0.0009)	(0.0004)			(0.0004)	(0.0003)			(0.0003)	(0.0003)			(0.0003)
Ln(Temperature)		-0.0002		0.0001		-0.0002		0.0003		0.0000		0.0001		-0.0002		0.0000		-0.0004		0.0000
		(0.0003)		(0.0003)		(0.0005)		(0.0007)		(0.0003)		(0.0004)		(0.0002)		(0.0003)		(0.0003)		(0.0004)
Ln(Precipitation)			0.0000	0.0000			-0.0001	-0.0001			0.0000	0.0000			-0.0001	-0.0001			-0.0001*	-0.0001
			(0.0000)	(0.0000)			(0.0001)	(0.0001)			(0.0000)	(0.0000)			(0.0000)	(0.0000)			(0.0000)	(0.0001)
Municip. × year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	94240	94240	88065	88065	94240	94240	88065	88065	94240	94240	88065	88065	94240	94240	88065	88065	94240	94240	88065	88065

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Robustness – Measurement

B.3.1 Main Mineral Price vs. Price Index

Table B.5: Price Index vs. Main Mineral Price

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price index)	0.0004*	0.0013**	0.0005**	0.0003	0.0002
	(0.0002)	(0.0005)	(0.0002)	(0.0002)	(0.0002)
Municip. $\times$ year FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Ln(Price index) denotes the natural logarithm of the weighted price index in month t. Weights are computed as each mineral's share of the *(i)* total production value or *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3.2 Outcomes

Table B.6: Alternative coding of outcomes

		Excl. months with protester riots				Excl. "unconfirmed" events	
	Protest	Protest	Arrests	Injuries	Casualties	Arrests	Injuries
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$M\times$ ln(Price)	0.0013	0.0012	0.0003*	0.0017**	0.0006**	0.0002	0.0011**
	(0.0011)	(0.0011)	(0.0002)	(0.0007)	(0.0002)	(0.0001)	(0.0005)
Municip. × year FEs	✓	✓	✓	✓	✓	✓	✓
Observations	94241	94210	94210	94210	94210	94224	94186

Notes: M equals one for mining municipalities (production or concessions), and 0 otherwise. Ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by *(i)* total production value and *(ii)* count of concession's primary commodities. Heteroskedasticity and autocorrelation corrected standard errors accounting for spatial correlation of up to 500km and unlimited serial correlation are obtained using the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Social Conflict vs. Event Location

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price)	0.0004	0.0010**	0.0003	0.0002	0.0002
	(0.0003)	(0.0005)	(0.0002)	(0.0002)	(0.0002)
Municip. × year FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3.3 Producing municipalities

Table B.8: Producing Municipalities

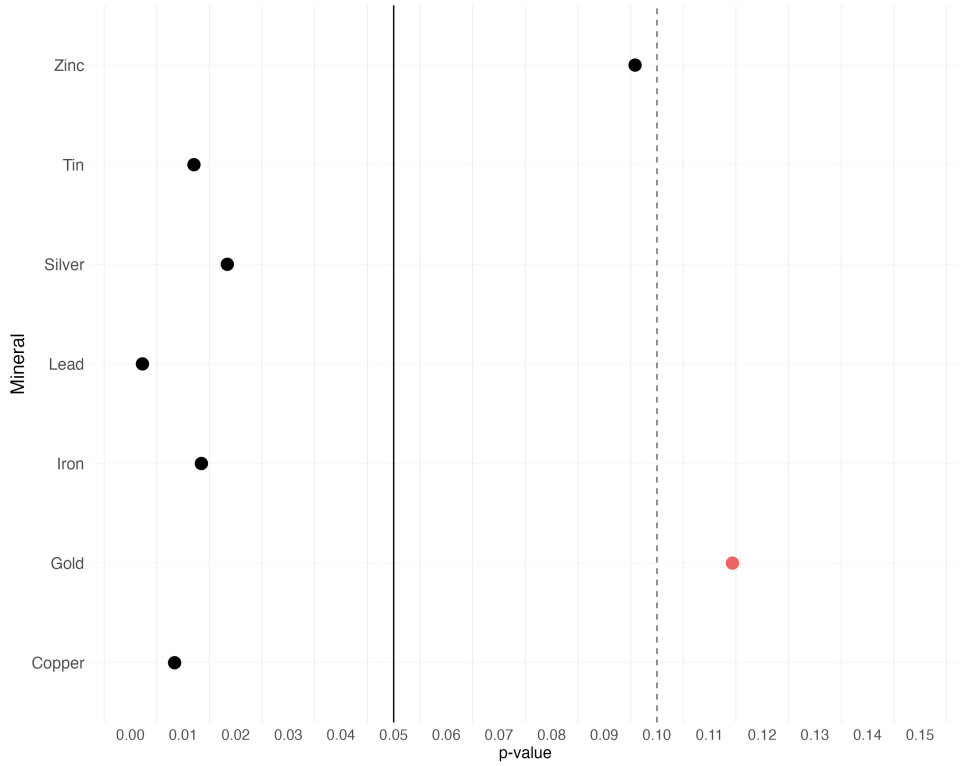
	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price)	0.0010*	0.0013	0.0006	0.0003	0.0001
	(0.0005)	(0.0009)	(0.0004)	(0.0002)	(0.0003)
Municip. $\times$ year FEs	✓	✓	✓	✓	✓
Observations	40850	40850	40850	40850	40850

Notes: M equals one for mining municipalities (production only) and 0 otherwise. Ln(Price) denotes the change in the logarithm of the main mineral price from month $t-1$ to t. The main mineral in a municipality is determined exclusively by the total production value. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.4 Robustness – Econometric Specification

B.4.1 Levels vs. Differences

Figure B.1: Unit Root Test



Notes: The p-values from Dickey–Fuller tests based on each mineral-specific *monthly* price series over the study period 2002–2019 are displayed. The null hypothesis is that the variable follows a random walk with nonzero drift. Price series have been purged of their common time components (i.e., I use the residuals from a regression of the log price on month × year dummies). The common certainty thresholds of 5% and 10% to reject the null hypothesis are depicted by the solid and dashed vertical lines, respectively.

Table B.9: First Difference – Main Mineral Price

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M \times \Delta$ ln(Price index)	0.0014	0.0028	0.0017*	-0.0007	-0.0005
	(0.0017)	(0.0021)	(0.0009)	(0.0013)	(0.0014)
Municip. × year FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Δ ln(Price) denotes the change in the natural logarithm of the main mineral price from month $t-1$ to t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: First Difference – Price Index

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M \times \Delta$ ln(Price index)	0.0014	0.0029	0.0020**	-0.0005	-0.0005
	(0.0017)	(0.0022)	(0.0010)	(0.0013)	(0.0015)
Municip. × year FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Δ ln(Price index) denotes the change in the natural logarithm of the weighted price index from month $t-1$ to t. Weights are computed as each mineral's share of the *(i)* total production value or *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Spatial Lags

	Force used against protesters						Protester behavior			
	Arrests		Injuries		Casualties		Violence		Riots	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$M\times$ ln(Price)	0.0002	0.0001	0.0024**	0.0020***	0.0008**	0.0006***	0.0000	0.0000	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0011)	(0.0006)	(0.0004)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
$M\times$ ln(Price neighbors [1st degree])	0.0006	0.0004	-0.0012	-0.0031**	-0.0003	-0.0010*	0.0005*	0.0002	0.0007**	0.0006
	(0.0004)	(0.0003)	(0.0014)	(0.0013)	(0.0005)	(0.0005)	(0.0002)	(0.0002)	(0.0003)	(0.0004)
$M\times$ ln(Price neighbors [2nd degree])		0.0004		0.0036***		0.0013***		0.0004		0.0002
		(0.0003)		(0.0010)		(0.0004)		(0.0003)		(0.0004)
Cumulative effect	0.0007**	0.0009**	0.0012*	0.0026***	0.0005	0.0010***	0.0005	0.0006*	0.0006*	0.0006
	(0.0004)	(0.0004)	(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Municip. $\times$ year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	93290	93290	93290	93290	93290	93290	93290	93290	93290	93290

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.12: Temporal Lags

	Force used against protesters									Protester behavior					
	Arrests			Injuries			Casualties			Violence			Riots		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
$M\times$ ln(Price $[t]$)	0.0014	0.0010	0.0014	0.0033	0.0034	0.0049	0.0019**	0.0020**	0.0036**	-0.0006	-0.0008	-0.0015	-0.0005	-0.0006	-0.0007
	(0.0017)	(0.0017)	(0.0025)	(0.0022)	(0.0023)	(0.0031)	(0.0009)	(0.0009)	(0.0017)	(0.0013)	(0.0013)	(0.0019)	(0.0014)	(0.0015)	(0.0020)
$M\times$ ln(Price $[t-1]$)	-0.0011	0.0016	0.0015	-0.0015	-0.0020	-0.0025	-0.0013	-0.0016	-0.0021	0.0009	0.0019	0.0020	0.0007	0.0013	0.0013
	(0.0017)	(0.0026)	(0.0027)	(0.0021)	(0.0031)	(0.0032)	(0.0009)	(0.0013)	(0.0014)	(0.0013)	(0.0019)	(0.0020)	(0.0015)	(0.0022)	(0.0023)
$M\times$ ln(Price $[t-2]$)		-0.0025	-0.0025		0.0005	0.0006		0.0003	0.0004		-0.0009	-0.0010		-0.0006	-0.0005
		(0.0016)	(0.0016)		(0.0018)	(0.0018)		(0.0008)	(0.0008)		(0.0008)	(0.0008)		(0.0010)	(0.0010)
$M\times$ ln(Price $[t+1]$)			-0.0003			-0.0012			-0.0014			0.0007			0.0001
			(0.0012)			(0.0017)			(0.0009)			(0.0007)			(0.0008)
Cumulative effect	0.0003*	0.0001	0.0001	0.0018**	0.0018**	0.0018**	0.0006**	0.0006**	0.0005*	0.0003*	0.0002*	0.0002**	0.0002	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0007)	(0.0008)	(0.0008)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)
Municip. × year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	94240	94240	93744	94240	94240	93744	94240	94240	93744	94240	94240	93744	94240	94240	93744

Notes: M equals one for mining municipalites (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.5 Robustness – Additional Sensitivity Checks

B.5.1 Seasonality

Table B.13: Month Fixed Effects

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
M × ln(Price)	0.0004*	0.0018**	0.0006**	0.0002	0.0001
	(0.0002)	(0.0007)	(0.0003)	(0.0002)	(0.0002)
District × year FEs	✓	✓	✓	✓	✓
Month FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions), and 0 otherwise. Ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by *(i)* total production value and *(ii)* count of concession's primary commodities. Heteroskedasticity and autocorrelation corrected standard errors accounting for spatial correlation of up to 500km and unlimited serial correlation are obtained using the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.5.2 World Market Share

Table B.14: Sample of Municipalities with Negligible World Market Share

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price)	0.0004**	0.0019**	0.0006**	0.0002	0.0002
	(0.0002)	(0.0007)	(0.0003)	(0.0002)	(0.0002)
Municip. $\times$ year FEs	✓	✓	✓	✓	✓
Observations	90060	90060	90060	90060	90060

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. Ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to 500 km and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.5.3 Spatial Kernel

Table B.15: Alternative Levels of Spatial Clustering

	Force used against protesters			Protester behavior	
	Arrests	Injuries	Casualties	Violence	Riots
	(1)	(2)	(3)	(4)	(5)
$M\times$ ln(Price)	0.0004	0.0019	0.0007	0.0002	0.0002
distance: 50	(0.0002)**	(0.0007)***	(0.0003)***	(0.0002)	(0.0002)
distance: 100	(0.0002)*	(0.0007)***	(0.0003)***	(0.0002)	(0.0002)
distance: 250	(0.0002)**	(0.0007)***	(0.0003)**	(0.0002)	(0.0002)
distance: 500	(0.0002)**	(0.0007)***	(0.0003)**	(0.0002)	(0.0002)
distance: 750	(0.0002)**	(0.0007)***	(0.0003)**	(0.0002)	(0.0002)
distance: 1000	(0.0002)**	(0.0007)***	(0.0003)**	(0.0002)	(0.0002)
Municip. × year FEs	✓	✓	✓	✓	✓
Observations	94240	94240	94240	94240	94240

Notes: M equals one for mining municipalities (production or concessions) and 0 otherwise. ln(Price) denotes the natural logarithm of the main mineral price in month t. The main mineral in a municipality is determined by the *(i)* total production value and *(ii)* count of a concession's primary commodities. Heteroskedasticity- and autocorrelation-corrected standard errors accounting for spatial correlation of up to the stated distance (in km) and unlimited serial correlation are obtained with the Stata module `acreg` (Colella et al., 2023). A linear decay in distance in the spatial correlation structure is assumed. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

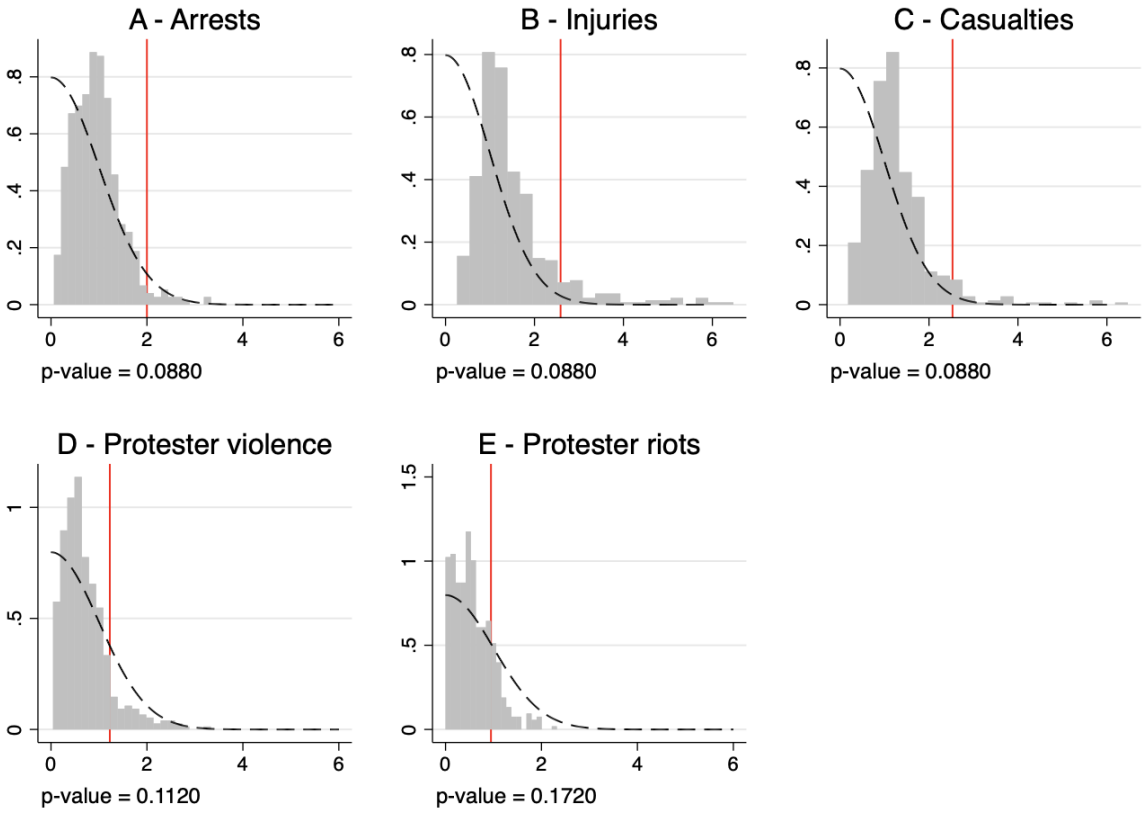
B.5.4 Multiple Hypothesis Correction

Table B.16: Romano–Wolf (Multiple Hypothesis Testing–Adjusted) P-Values

	Outcome	Model p-value	Resampled p-value	Romano-Wolf p-value
Force used against protesters	Arrests	0.045	0.008	0.088
	Injuries	0.010	0.044	0.088
	Casualties	0.011	0.074	0.088
Protester behavior	Violence	0.220	0.172	0.172
	Riots	0.343	0.100	0.112

Notes: The Romano–Wolf p-values adjusted for multiple hypothesis testing are calculated with the resampled null distribution from 500 bootstrap samples with the *Stata* command `rwolf` (Clarke et al., 2020).

Figure B.2: Null Distributions and Original T Statistics

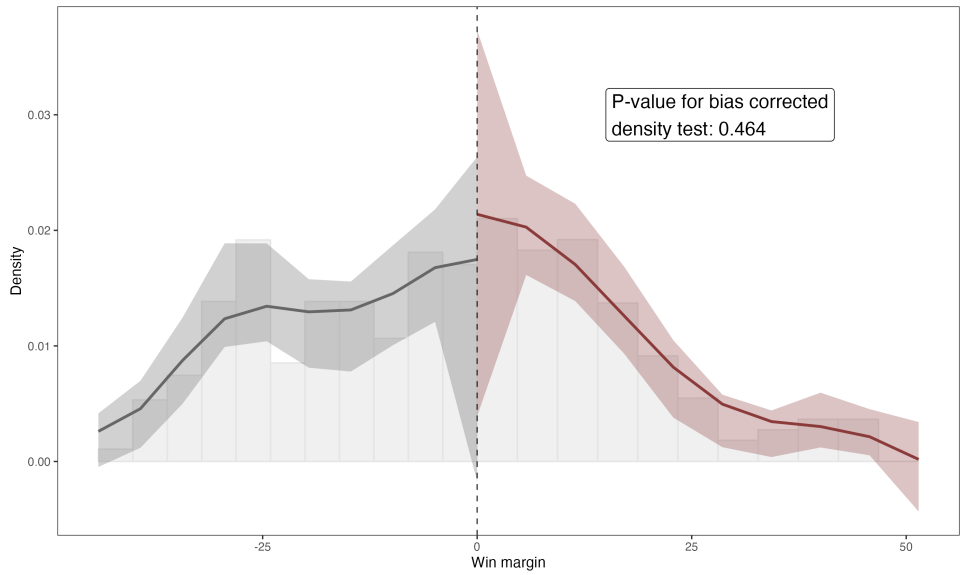


Notes: Each panel documents the null distributions used for the calculation of the *Romano–Wolf* adjusted p-values for each of the 5 baseline dependent variables. The Romano–Wolf adjusted p-values are calculated with the *Stata* command `rwolf` (Clarke et al., 2020) and displayed below each panel. The histogram in each panel depicts the stepdown resampled null distribution from 500 bootstrap samples. The dashed line captures the theoretical half-normal, and the solid vertical line presents the original t statistic corresponding to each outcome.

C Pro-Mining Local Politicians, State Repression, and Rent Seeking

C.1 Manipulation Test

Figure C.1: Manipulation Test – Pro- vs. Anti-Mining Candidates



Notes: This figure presents the manipulation test suggested by Cattaneo et al. (2018) and implemented in the R/Stata command `rddensity` using a quadratic polynomial and triangular kernel weights. The p-value for the bias-corrected density test is 0.464.

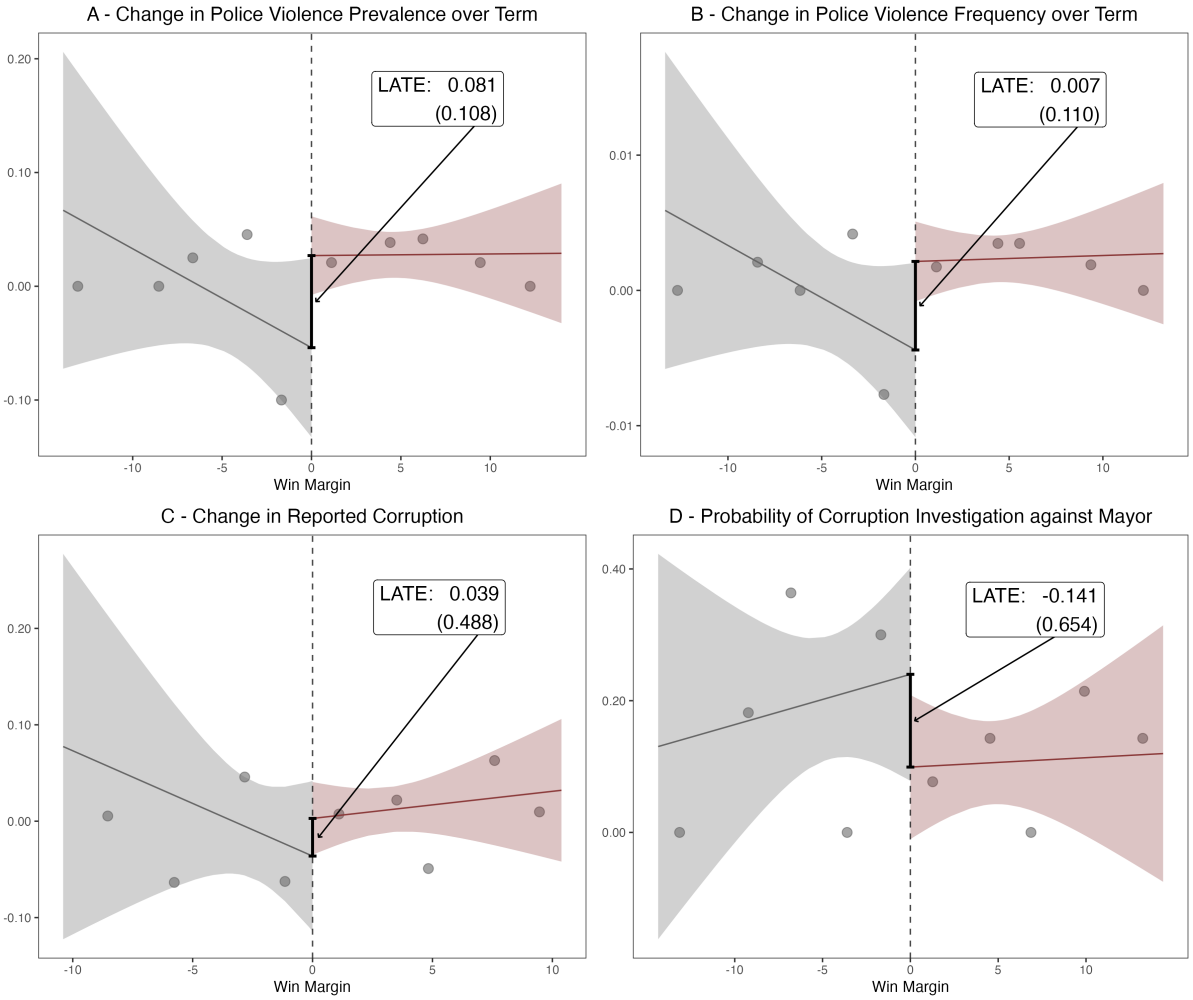
C.2 Additonal RDD Results

Table C.1: Smooth Covariates near the Cut-off

	RD Estimate	Robust p-value	Robust 90% CI	Band-width	Mean	Obs.	Effect. Obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Political competition	0.097	0.917	[-0.98, 1.11]	10.0	4.795	234	88
Turnout	4.630	0.142	[-0.667, 11.7]	8.8	82.571	234	78
Canon minero	-289.000	0.635	[-587, 1063]	9.6	959.916	230	83
Royalties	-1.060	0.845	[-475, 603]	6.9	315.357	230	61
Area	3.790	0.962	[-511, 541]	10.6	703.761	234	93
Nighttime lights	0.314	0.823	[-1.57, 2.07]	10.4	2.270	230	89
River length	-120.000	0.439	[-384, 138]	11.3	211.121	234	97
Elevation	-71.900	0.478	[-849, 337]	11.6	3423.230	234	98
Lake area share	0.003	0.393	[-0.004, 0.012]	11.1	0.004	234	96
Indigenous (0/1)	-0.095	0.250	[-0.208, 0.037]	17.3	0.021	234	136
Population density	-5.630	0.981	[-43.9, 45.2]	10.3	28.988	234	90
Road density	0.006	0.861	[-0.043, 0.054]	11.1	0.058	234	96
Provincial capital (0/1)	-0.038	0.675	[-0.419, 0.249]	12.9	0.184	234	107
ENAHO Respondents	9.290	0.941	[-46.2, 50.5]	9.0	102.271	85	27

Notes: Local linear regression estimates from Calonico et al. (2014) with triangular kernel weights and optimal MSE bandwidth are reported in column 1. Columns 2 to 7 report 90% robust confidence intervals and p-values adjusted for clustering at the regional level, the optimal MSE bandwidth, the number of observations in the sample and in the bandwidth, and the dependent variable mean. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.2: Effect of Electing a Pro-mining Mayoral Candidate on Police Violence



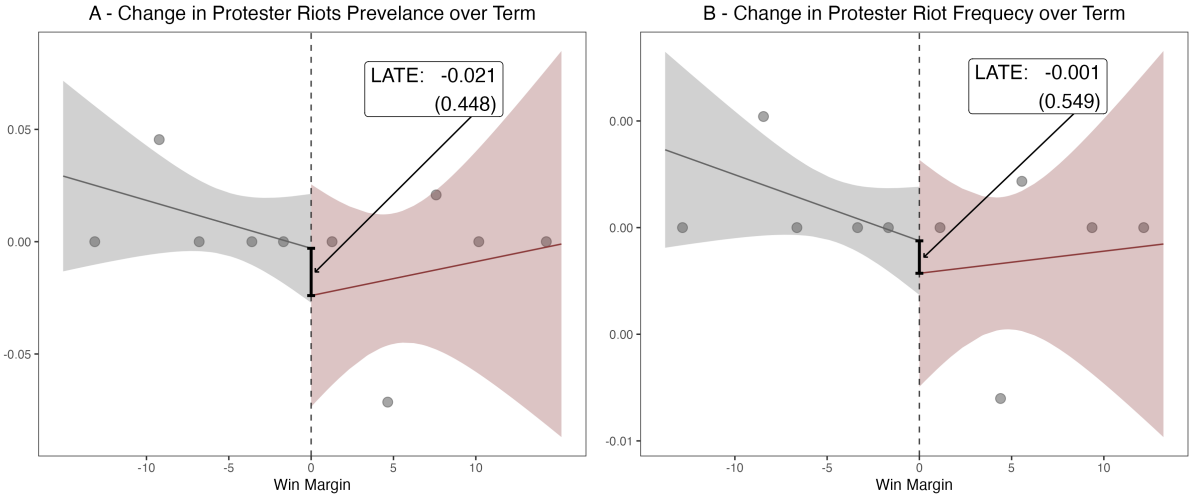
Notes: This figure presents a graphical approximation of the regression discontinuity design. Each panel presents the local average treatment effect (LATE) of (narrowly) electing a pro-mining politician over an anti-mining politician. The LATE point estimate and its robust p-value adjusted for clustering at the regional level are reported. Shaded areas denote 90% confidence intervals. Observations are shown within MSE-optimal bandwidths.

Table C.2: Protester Riots

	Protester Riot Prevalence		Protester Riot Frequency	
	(1)	(2)	(3)	(4)
Pro-Mining Mayor	-0.021	-0.024	-0.001	-0.002
Cluster-robust p-value	0.448	0.224	0.549	0.189
90% CI	[-0.058, 0.026]	[-0.064, 0.015]	[-0.004, 0.002]	[-0.004, 0.001]
Baseline Covariates		✓		✓
Bandwidth	15.7	15.9	13.9	12.6
Dep. variable mean	0.001	0.001	0.000	0.000
Observations	230	230	230	230
Effective obs.	124	124	112	102

Notes: Local linear regression estimates from Calonico et al. (2014) with triangular kernel weights and optimal MSE bandwidth are reported. Even columns present local covariate-adjusted linear estimates. The "initial investigation" sample excludes mayors that faced legal persecution. 90% robust confidence intervals and p-values adjusted for clustering at the regional level, the number of observations in the sample and in the bandwidth, the optimal MSE bandwidth, and the dependent variable mean are reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.3: Effect of Electing a Pro-mining Mayoral Candidate on Protester Riots



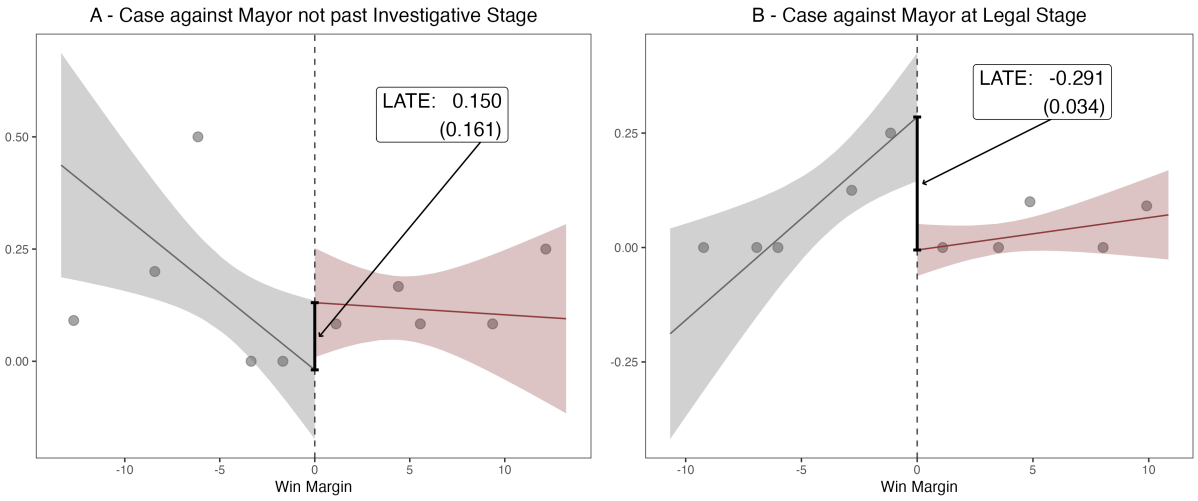
Notes: This figure presents a graphical approximation of the regression discontinuity design. Each panel presents the local average treatment effect (LATE) of (narrowly) electing a pro-mining politician over an anti-mining politician. The LATE point estimate and its robust p-value adjusted for clustering at the regional level are reported. Shaded areas denote 90% confidence intervals. Observations are shown within MSE-optimal bandwidths.

Table C.3: Stage of Criminal Process against Former Mayor

	Initial Investigation		Legal Persecution	
	(1)	(2)	(3)	(4)
Pro-Mining Mayor	0.163	0.193	-0.291**	-0.264**
Cluster-robust p-value	0.134	0.211	0.029	0.04
90% CI	[-0.02, 0.44]	[-0.071, 0.524]	[-0.629, -0.089]	[-0.579, -0.065]
Baseline Covariates		✓		✓
Bandwidth	13.3	8.4	10.9	11.3
Dep. variable mean	0.110	0.110	0.064	0.064
Observations	219	217	234	230
Effective obs.	106	69	95	95

Notes: Local linear regression estimates from Calonico et al. (2014) with triangular kernel weights and optimal MSE bandwidth are reported. Even columns present local covariate-adjusted linear estimates. The "initial investigation" sample excludes mayors that faced legal persecution. 90% robust confidence intervals and p-values adjusted for clustering at the regional level, the number of observations in the sample and in the bandwidth, the optimal MSE bandwidth, and the dependent variable mean are reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.4: Stage of Criminal Process against Former Mayor



Notes: This figure presents a graphical approximation of the regression discontinuity design. Each panel presents the local average treatment effect (LATE) of (narrowly) electing a pro-mining politician over an anti-mining politician. The LATE point estimate and its robust p-value adjusted for clustering at the regional level are reported. Shaded areas denote 90% confidence intervals. Observations are shown within MSE-optimal bandwidths.

D Democratic Accountability

D.1 Additional Results

Table D.1: Separate RD Equations for the Effect on Electoral Performance of Incumbents

	Unconditional									Conditional								
	Elected			Vote Share			Runs			Elected			Vote Share			Margin		
	RDD (1)	RDD (2)	Diff (3)	RDD (4)	RDD (5)	Diff (6)	RDD (7)	RDD (8)	Diff (9)	RDD (10)	RDD (11)	Diff (12)	RDD (13)	RDD (14)	Diff (15)	RDD (16)	RDD (17)	Diff (18)
Incumbency	-25.818***	-2.697	23.121	-6.201***	-0.196	6.005	-16.859**	5.173	22.033	-31.275***	1.698	32.973	-0.071	-0.203	-0.132	-6.916***	-9.750	-2.834
	(0.000)	(0.647)	(0.296)	(0.002)	(0.834)	(0.461)	(0.013)	(0.974)	(0.389)	(0.000)	(0.920)	(0.263)	(0.863)	(0.808)	(0.951)	(0.001)	(0.130)	(0.749)
Police Violence		✓			✓			✓			✓			✓			✓	
Bandwidth	8.8	9.1		7.8	12.6		8.9	9.6		8.0	11.5		8.9	7.4		9.4	7.9	
Dep. var. mean (%)	23.088	18.085		16.253	14.509		62.912	59.574		36.698	30.357		28.424	26.917		-4.423	-4.483	
Observations	2850	94		2850	94		2850	94		1793	56		1793	56		1793	56	
Eff. Obs.	1736	52		1596	68		1752	52		1026	42		1112	32		1157	35	

Notes: Local linear regression estimates with triangular kernel weights and optimal MSE bandwidth are reported. Robust p-values adjusted for clustering at the regional level in parentheses. The number of observations in the sample and the bandwidth, the optimal MSE bandwidth, and the dependent variable mean (in percent) are reported. P-values for the difference in RD estimates (column denoted "Diff") are calculated using a two-tailed Z-test (Paternoster et al., 1998). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E Is Violent State Repression Effective?

E.1 Descriptive Statistics

Table E.1: Descriptive Statistics – Social Conflict Level

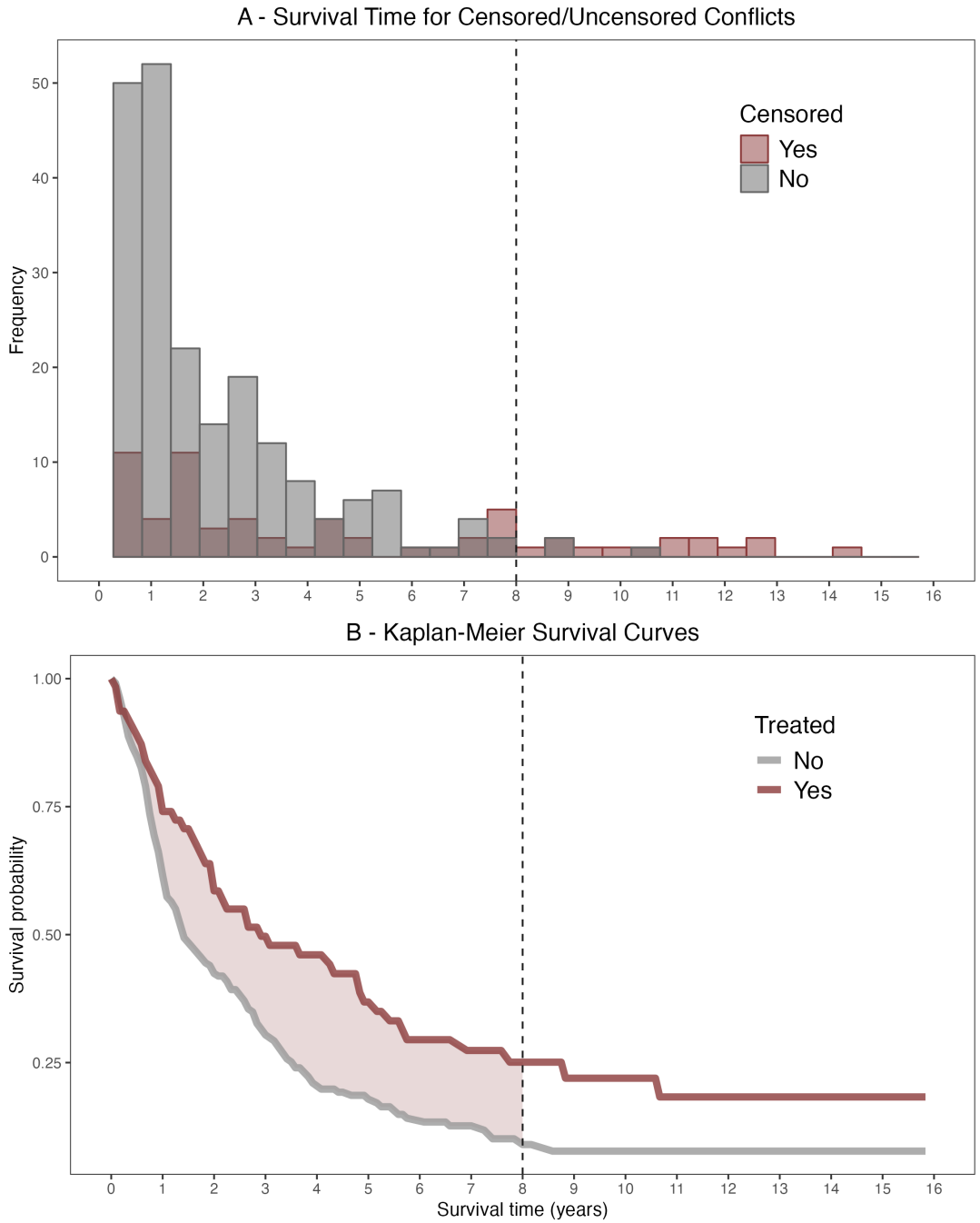
	Resolution (N=138)				Removed (N=84)				Right-censored (N=76)			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Time-varying Covariates												
Δ Main mineral price (in %)	7.22	28.10	-57.91	135.84	-0.80	10.36	-22.37	52.97	2.54	9.63	-16.30	40.52
Canon minero (in million USD)	2.83	6.38	0.00	43.45	2.24	6.56	0.00	46.21	3.64	8.78	0.00	47.77
Nighttime lights	2.68	4.23	0.00	35.27	1.98	2.62	0.00	12.12	4.70	4.20	0.03	27.71
No. of months w/ protester riots	0.02	0.11	0.00	1.00	0.02	0.12	0.00	1.00	0.08	0.27	0.00	2.00
No. of months w/ protests	0.53	0.58	0.00	3.00	0.58	0.58	0.00	3.00	0.77	0.86	0.00	4.29
Royalties (in million USD)	0.64	2.37	0.00	25.62	0.33	1.15	0.00	7.39	1.27	3.51	0.00	21.06
Share of conflict-years w/ police violence (in %)	10.98	28.13	0.00	100.00	12.05	27.81	0.00	100.00	8.18	19.75	0.00	100.00
Time-invariant Baseline Covariates												
Area (in km^2)	713.96	749.53	45.62	5765.65	512.53	452.81	39.17	2927.89	783.05	982.79	15.64	5765.65
Duration (in years)	1.99	1.53	1.00	9.00	2.88	2.02	1.00	9.00	5.00	4.12	1.00	16.00
Elevation (in km)	3.46	1.20	0.04	4.75	3.71	0.80	0.75	4.65	3.72	0.89	0.04	4.69
Foreign ownership (0/1)	0.54	0.50	0.00	1.00	0.43	0.50	0.00	1.00	0.50	0.50	0.00	1.00
Indigenous communities (0/1)	0.01	0.12	0.00	1.00	0.01	0.11	0.00	1.00	0.01	0.11	0.00	1.00
Market value of majority owner (in billion USD)	7.16	21.86	0.00	105.56	6.83	22.24	0.00	97.78	9.59	22.72	0.00	88.48
Population density	52.45	132.41	1.78	1224.96	27.22	32.31	0.89	195.43	50.09	203.59	0.95	1778.34
River density	0.26	0.05	0.10	0.40	0.25	0.04	0.17	0.37	0.26	0.05	0.11	0.37
Road density	0.06	0.06	0.00	0.22	0.06	0.06	0.00	0.24	0.05	0.06	0.00	0.24
Share of area covered by Lakes (in %)	0.43	1.03	0.00	5.59	0.43	1.63	0.00	14.53	0.57	1.70	0.00	13.39

Notes: The unit of observation is the social conflict level and Time-varying variables are measured at *conflict-year* intervals. The share of conflict-years with police violence captures in how many of the total conflict-years at least one incidence of police violence against protesters was reported.

E.2 State Repression and Social Conflict Duration

Using the framework of a survival analysis, I examine if violent state repression prolongs the duration of social conflicts.

Figure E.1: Social Conflict Duration



Notes: Panel A presents of histogram of outcomes by event status (observed vs. censored) and the chosen truncation time of 8 years (vertical dashed line). Lines in Panel B depict the Kaplan-Meier survival curves for the treatment (red) and control (gray) group. The red shaded are in Panel B displays the difference in the unadjusted restricted mean survival times (RMST) for both groups.

RMST The goal is to measure the average effect of treatment A_i on survival time T_i conditional on set of predetermined covariates X_i.[46] Here, treatment is a binary indicator if police violence against protesters has ever been observed in relation to social conflict i. Following Cui et al. (2023), I estimate the treatment effect on the *restricted mean survival time (RMST)* for a pre-defined maximal time horizon h. Since the longest lasting social conflicts are all censored (s. Panel A in Figure E.1), this strategy allows me to treat any conflict tracked past the horizon h as "observed" (even if they are eventually censored), ensuring that the moments of T_i are identified. For the remainder of the analysis I set h to 8 years, which is just below the maximum duration of treated social conflicts of about 9 and a half years.

Table E.2: Effect of Police Violence on Social Conflict Duration

	RMST	RMST	2SKM	CSF
	(1)	(2)	(3)	(4)
Police violence	1.340***	1.307***	1.250**	1.318***
95% CI	[0.507,2.172]	[0.466,2.148]	[0.370,2.089]	[0.418,2.217]
Covariates		✓	✓	✓
Observations	298	298	298	298

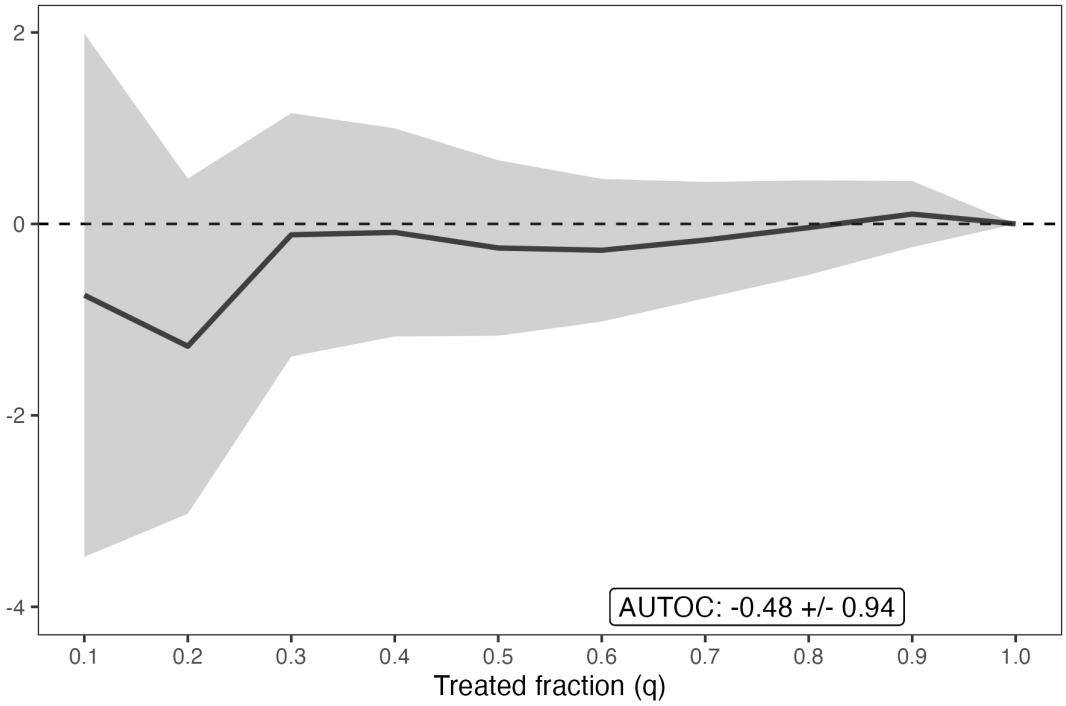
Notes: Column 1 and 2 present the difference in unadjusted, respectively, adjusted restricted mean survival times (RMST) for social conflicts that experienced at least one incident of police violence vs. not using the method proposed by Tian et al. (2013) and Uno et al. (2014). Column 3 presents estimates for the average treatment effect (ATE) using the 2SKM method proposed by Sant'Anna (2016). Column 4 presents the ATE estimate using causal survival forests as implemented in Cui et al. (2023). Estimators in columns 2 and 3 control only for a subset of pre-determined covariates (i.e. indicator for indigenous communities, population density, canon minero, and royalties), whereas the full set of predetermined covariates is used by the causal survival forests analysis in column 4. 95% confidence intervals are presented. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results Table E.2 presents the estimates for the effect of violent protest suppression on social conflict duration. Column 1 presents the difference in the unadjusted RMST between treatment and control group, which corresponds simply to the difference in the area under the curve of the survival function up to horizon h (s. Panel B in Figure E.1 for a graphical illustration), whereas Column 2 reports the estimated difference in RMSTs after adjusting for a subset of the predetermined covariates using the method proposed by Tian et al. (2013) and Uno et al. (2014). In specific, convergence requires that the covariate

[46] Note that the full set of pre-determined covariates here includes the time-invariant baseline covariates listed in Table E.1 and the time-varying covariates measured at the start of the conflict.

set is restricted to an indicator for indigenous communities, population density, canon minero and royalty transfers. Estimates for both methods suggest that the use of violence prolongs the RMST by about 1.3 years. The use of Sant'Anna (2016)'s *two-step Kaplan-Meier estimator (2SKM)* and of (Cui et al., 2023)'s *causal survival forests (CSF)* confirms the sizable positive and significant average treatment effect (ATE) on the RMST. The flexible non-parametric CSF estimator allows the inclusion of all predetermined covariates and does not require censoring to be independent of baseline covariates after conditioning on duration and treatment status, as it is the case for 2SKM. Finally, formal tests reveal no meaningful heterogeneity of treatment effects across conflicts (Figure E.2).

Figure E.2: Treatment Effect Heterogeneity



Notes: The *Targeting Operator Characteristic (TOC)* curve is presented. It depicts the difference between the average treatment effect (ATE) of the q-th fraction of conflicts with the largest estimated increase in conflict duration and the overall ATE. Shaded areas denote 95% confidence intervals. In addition, the area under the TOC curve ($AUTOC$) and its 95% confidence interval is presented.

E.3 Dynamic Causal Inference and Incremental Causal Effects

E.3.1 Technical Appendix

In this section, I give a brief overview on *(i)* why traditional regression methods in the context of a dynamic treatment process fail to recover the causal effect of interest, *(ii)* how dynamic causal inference methods such as *marginal structural models (MSM)* and *incremental causal effects* are able to adjust for time-varying confounders that are themselves affected by treatment, and *(iii)* which identifying assumptions have to hold. For more details on MSM and incremental causal effects, the interested reader is referred to Blackwell (2013) and Bonvini et al. (2023), respectively.

Set-Up For each social conflict, we observe:

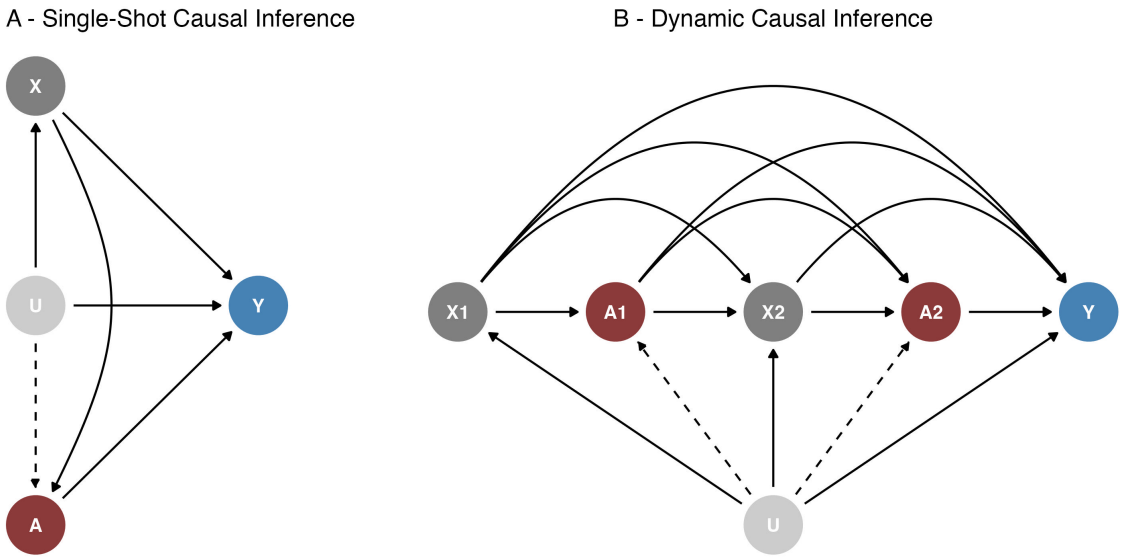
$$\mathbf{Z} = (\mathbf{X}_1, A_1, R_2, R_2(\mathbf{X}_2, A_2), ..., R_T, R_T(\mathbf{X}_{t-1}, A_{t-1}), R_{t+1}, R_{t+1}Y),$$

where $\mathbf{X}_t$ is a set of potentially time-varying covariates, A_t is a binary treatment indicator equaling one if at least one incident of police violence against protesters is registered at time t, R_t is a binary indicator equaling one if the conflict is still observed at time t, and Y denotes the binary outcome at the end of the conflict ($t = T$) that equals one if the conflict ends with a resolution agreement. Following the convention in the literature, let overbars denote past values of a variable such that, for instance, $\overline{\mathbf{X}}_t = (\mathbf{X}_1, ..., \mathbf{X}_t)$ and $\overline{A}_t = (A_1, ..., A_t)$, and, for simplicity, let $\mathbf{H}_t = \left(\overline{\mathbf{X}}_t, \overline{A}_{t-1}\right)$ denote all history just prior to treatment at time t. Finally, let lower-case letters a_t, $\mathbf{x}_t$, $\mathbf{h}_t$ represent realized values for A_t, $\mathbf{X}_t$, and $\mathbf{H}_t$. For the remainder of this section, I abstract from right-censoring and refer to Kim et al. (2021) for more details on the identification of incremental causal effects under censoring.

Why Traditional Methods Fail The decision of local authorities to employ violence against protesters can be modeled as a dynamic decision process that takes into account past developments such as the number of protests and riots or the evolution of mining revenues. Traditional regression methods, however, have to assume that all decisions are made at a single point in order to estimate the effect of violence on the social conflict outcome. The causal relationship between observed covariates ($\mathbf{X}$), unobserved confounders (U), treatment (A), and the outcome (Y) for this "single-shot" causal inference approach (Blackwell, 2013) is graphically illustrated in the directed acyclic graph (DAG) in Panel A of Figure E.3. For the effect of interest to be correctly identified the assumption of *no omitted variables* needs to hold (i.e. the dotted arrow in Panel A of Figure E.3 has to be absent). In the context of this study we would, therefore, like to control for, e.g., the number of protests associated with the social conflict, as both, the conflict outcome

and the decision to crack down on protesters, are affected by the (observed) strength of the opposition to the mining project. At the same time, protest behavior is affected by earlier incidents of police violence due to, e.g., intimidation or retaliation. The number of protests is, hence, an example of a *time-varying* confounder that both impacts the treatment decision and is affected by the treatment. In this scenario causal identification with traditional regression methods is not possible: including the confounder in the estimation equation introduces *posttreatment bias* while omitting the confounder will create *omitted variable bias*.

Figure E.3: Directed Acyclic Graphs (DAGs)



Notes: Y denotes the outcome, A the treatment, and X denotes (any and all) *observed* confounders. In Panel B, treatment and covariates at time $t \in \{1, 2\}$ are depicted. U represents (any and all) *unobserved* variables. Each arrow represents a causal relationships. Assuming *no omitted variables* in Panel A and *exchangability* in Panel B means assuming that the *dotted* arrows are not present.

Dynamic Causal Inference In contrast to standard regression models, dynamic causal inference methods such as MSM and incremental interventions can adjust for time-varying (observed) confounders while avoiding posttreatment bias. Consider the causal structure presented of the DAG in Panel B of Figure E.3. At each timepoint $t > 1$, the dynamic causal model will account for the impact of past values of both treatment and confounders on the decision to employ excessive force against protesters. E.g. treatment at $t = 2$ ($A2$) depends on the value of confounders $X2$ which are themselves allowed to dependent on their previous values ($X1$) and the treatment decision in the earlier period ($A1$).

Let $Y^{\overline{a}_T}$ denote the *potential outcome* that would have been observed under the treatment sequence $\overline{a}_T = (a_1, ...a_T)$. In contrast to the "single-shot" approach with only two

potential outcomes, i.e. $E\left[Y^{a=1}\right]$ and $E\left[Y^{a=0}\right]$, there exists 2^T potential outcomes in the dynamic causal inference framework, one for each possible treatment sequence. For the two period case in Panel B of Figure E.3, there exist $2^T = 2^2$ possible treatment sequences, i.e.: $\{0,0\}$, $\{0,1\}$, $\{1,0\}$ and $\{1,1\}$. For the case of $T = 16$ (as in the context of this study), there exist $2^{16} = 65,536$ potential outcomes.

MSM vs. Incremental Causal Effects To overcome this curse of dimensionality, MSM models impose a parametric structure on the model for the mean of the potential outcome, e.g.:

$$E\left[Y^{\overline{a}_T}\right] = g\left(\overline{a}_T, \delta\right) = \delta_0 + \delta_1 \sum_{t=1}^{T} a_t$$

. This model assumes that subjects (here conflicts) with the same number of treatment periods have similar potential outcomes, i.e., the timing of the treatment is assumed to have no impact on the final outcome. For instance, in the five period case ($T = 5$), the model assumes that the treatment sequence $\overline{a}_5 = \{a_1, ..., a_5\} = \{1,0,0,0,1\}$ has the same potential outcome as the treatment sequence $\{0,0,0,1,1\}$. The parameters of the MSM are then usually estimated using inverse probability weighting (IPW). Specifically, for each conflict a weight is constructed as the inverse product of conditional probabilities of receiving the observed treatment sequence, i.e.:

$$W = \prod_{t=1}^{T} \frac{1}{\left(A_t = a_t \mid \overline{\mathbf{X}}_t = \overline{\mathbf{x}}_t, \overline{A}_{t-1} = \overline{a}_{t-1}\right)}.$$

The denominator can be estimated using the *time-dependent* propensity score model $\mathbb{P}\left(A_t \mid \overline{\mathbf{H}}_t\right) = \pi_t(\mathbf{h}_t)$, with a common choice being a logistic regression model estimated separately at each time period. If propensity score estimates are extreme (close to 0 or 1), the IPWs become very unstable. A critical identification assumption of MSMs, therefore, is *positivity*, i.e. that the propensity scores are bounded away from 0 and 1. In other words, MSMs require that each subject has a nonzero chance of receiving any possible treatment sequence at every time point. In practice, this assumption is unlikely to be fulfilled, particularly in studies with many time points. Importantly, even if *positivity* by design holds as in the case of a randomized experiment, only a small fraction of possible treatment regimes will be observed because the number of possible treatment sequences grows exponentially with the number of time points (curse of dimensionality).

Incremental interventions, on the other hand, avoid *positivity* requirements altogether by shifting the propensity score π_t by δ instead of setting A_t. The causal estimand of interest is the average counterfactual outcome (across social conflicts) if the odds of treatment were multiplied by δ, i.e. $\psi\left(\delta\right) = \mathbb{E}\left[Y^{\mathbf{Q}(\delta)}\right]$, where $Y^{\mathbf{Q}(\delta)}$ is the potential

outcome under the treatment process $(q_1, ..., q_T)$ and:

$$q_t(\mathbf{h}_t, \delta, \pi_t) = \frac{\delta \pi_t(\mathbf{h}_t)}{\delta \pi_t(\mathbf{h}_t) + 1 - \pi_t(\mathbf{h}_t)}.$$

This estimation strategy has three important implications. First, the *incremental effect* $\psi(\delta)$ is a compounding effect of two different changes: (1) multiplying π_t by δ and (2) covariates at earlier timepoints changing due to the incremental intervention. Second, *incremental causal effects* do not require *positivity*. To see why, let's consider the case where $T = 1$. In this case there are two sets of conflicts for which the *positivity* assumption is violated: conflicts that are never treated ($\pi = 0$) and conflicts that always experience police violence ($\pi = 1$). For these two groups, incremental interventions do not alter the propensity score, i.e. $\pi = 0 \Rightarrow q = 0$ and $\pi = 1 \Rightarrow q = 1$. Third, this estimator is completely *nonparametric*.

Independent of the chosen method (MSM or *incremental causal effects*) the following two (untestable) identification assumptions have to be fulfilled:

1. *Consistency*: $Y = Y^{\overline{a}_t}$ if $\overline{A}_t = \overline{a}_t$
2. *Exchangability*: $A_t \perp\!\!\!\perp Y^{\overline{a}_t} \mid \overline{\mathbf{H}}_t$

Consistency requires that the observed outcomes equal the corresponding potential outcomes under the observed treatment process. This assumption would be violated in the case of interference, e.g., if police violence in one conflict has an effect on the final outcome of another social conflict. *Exchangability* assumes that treatment each period is as good as random conditional on the past. In other words, *exchangability* demands that there are no unmeasured confounders (i.e. that the dotted arrows in Panel B of Figure E.3 are not present).

E.3.2 Additional Results

Table E.3: State Repression and Resolution Agreement Probability

	Resolution Aggrement Probability		Difference in Resolution Agreement Probability	
δ	Prob	90% CI	Diff.	90% CI
0.1	76.67	[59.20, 94.14]	17.07	[1.85, 32.29]
0.3	69.86	[59.22, 80.51]	10.26	[3.01, 17.52]
0.5	65.41	[57.00, 73.82]	5.81	[1.52, 10.10]
0.7	63.19	[55.75, 70.64]	3.59	[0.91, 6.27]
1.0	59.60	[52.27, 66.93]	—	—
1.5	54.81	[43.04, 66.59]	-4.79	[-12.08, 2.50]
2.0	45.17	[22.79, 67.56]	-14.43	[-33.35, 4.49]
2.5	33.21	[-3.23, 69.64]	-26.40	[-59.75, 6.96]
3.0	18.90	[-34.25, 72.05]	-40.71	[-90.95, 9.54]

Notes: Column 1 presents the estimated resolution agreement probability $\psi(\delta)$ if the police violence odds were multiplied by factor δ. Column 3 displays the estimated difference in resolution probability relative to no intervention ($\delta = 1$). The corresponding pointwise 90% confidence interval is reported in column 2, respectively, 4.

Appendix References

Cardellino, C. (2019). Spanish Billion Words Corpus and Embeddings.

Clarke, D., Romano, J. P., and Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in Stata. *Stata Journal*, 20(4):812–843.

Colella, F., Lalive, R., Sakalli, S. O., and Thoenig, M. (2023). acreg: Arbitrary correlation regression. *Stata Journal*, 23(1):119–147.

Cui, Y., Kosorok, M. R., Sverdrup, E., Wager, S., and Zhu, R. (2023). Estimating heterogeneous treatment effects with right-censored data via causal survival forests. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(2):179–211.

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific data*, 7(1):109.

Honnibal, M., Montani, I., Van Landeghem, S., and Boyd, A. (2020). spaCy: Industrial-strength Natural Language Processing in Python.

Monfreda, C., Ramankutty, N., and Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22(1).

Paternoster, R., Brame, R., Mazerolle, P., and Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36(4):859–866.

Sant'Anna, P. H. (2016). Program evaluation with right-censored data. *arXiv preprint arXiv:1604.02642*.

Tian, L., Zhao, L., and Wei, L. J. (2013). Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. *Biostatistics*, 15(2):222–233.

Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., Schrag, D., Takeuchi, M., Uyama, Y., Zhao, L., Skali, H., Solomon, S., Jacobus, S., Hughes, M., Packer, M., and Wei, L.-J. (2014). Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology*, 32(22):2380–2385.