Inequality and Social Unrest in India

Mathieu Couttenier
Ecole Normale Superieure
15 Parvis Descares
69007 Lyon
France
mail : mathieu.couttenier@ens-lyon.fr

Jérémy Laurent-Lucchetti
University of Geneva
Institute of Economics and Econometrics, GSEM
40 Boulevard du pont d'arve
1211 Geneva
Switzerland
Office: +41 22 379 82 81
Mail: jeremy.lucchetti@unige.ch

Lore Vandewalle
The Geneva Graduate Institute
Department of Economics
Maison de la paix
Chemin Eugène-Rigot 2
1202 Genève
Switzerland
Office: +41 22 908 59 21
Mail : lore.vandewalle@graduateinstitute.ch

Inequality and Social Unrest in India

Mathieu Couttenier[*], Jeremy Laurent-Lucchetti[†] and Lore Vandewalle[‡]

Tuesday 22nd November, 2022

Abstract

We show that inequality triggers social unrest in rural India. We develop a theoretical framework where social unrest is rationally used by civilians to oppose (unfair) surplus sharing by the elite. We predict that the probability of observing social unrest in a village increases with the sum of distances between the (log) average and the lowest incomes. We bring our measure to the data using bank account details in 2,197 Indian villages. We document that a 10% increase in our inequality measure increases by 6.5% the unconditional probability of observing social unrest in a village in a given month.

JEL Codes: D31, D63, D74
Keywords: Inequality, Social unrest, Development, Redistribution.

[*]ENS de Lyon & CEPR, France. Email: mathieu.couttenier@ens-lyon.fr.
[†]GSEM, Geneva, Switzerland. Email: Jeremy.Lucchetti@unige.ch.
[‡]The Geneva Graduate Institute, Switzerland. Email: lore.vandewalle@graduateinstitute.ch.
We are grateful to Sanjay Prasad for excellent research assistance and to the staff at Basix Sub-K for their continued support and collaboration. We specifically thank Debraj Ray for insightful comments on a previous version of the paper. All remaining errors are of course our own.

1 Introduction

Inequality and social unrest are on the rise. In 2019, around 60% of countries in the world experienced violent demonstrations and since 2011, the overall level has increased by 282% (IEP, 2020). This surge in social unrest swept through all continents, which is unprecedented since the wave of movements in the late 1980s in Eastern Europe and Asia. Economic inequalities are usually put forward as a crucial motivation of this surge in violence,[1] along with perceived corruption of governing elite, rebellion against authoritarianism and general political discontent. International organizations, such as the IMF, OECD and World Bank, mention that this is a threat for political and macroeconomic stability as inequalities and unrest "complicate sound economic policy making".[2]

Surprisingly, empirical regularities linking inequality and social unrest, and explaining the emergence of social unrest within a country are limited (Ray and Esteban, 2017). Three main reasons come to the fore. First, most existing papers have estimated pooled cross-country regressions (Muller and Seligson, 1987; Nafziger and Auvinen, 2002). A concern is that countries with different levels of inequality may differ on other (unobserved) dimensions. This makes the causal interpretation and the identification of the mechanisms not straightforward. Second, most of the existing studies link inequality to large political conflicts, i.e. civil wars. These events usually require substantial resources and organizational capacity, dampened by poverty and excessive inequality. Social unrest, on the other hand, is local, spontaneous and requires very few resources. As stated by Ray and Esteban (2017) *"it is plausible that the dominant form taken by the class struggle envisioned by Marx is social unrest—strikes, demonstrations, etc.—rather than armed civil war."* To identify the effect on local social unrest, one needs extensive data on inequality *within countries* and across time. It is challenging to find good sources of data though, in particular for low- and middle income countries. Aggregate measures at the country level—like in most of the existing literature—tend to mask, rather than reflect, the inherent diversity in local inequalities that might trigger events of social unrest. Finally, the lack of empirical evidence might also be driven by the absence of a theoretically relevant measure of inequality to guide an empirical estimation. The theoretical literature linking inequalities and conflict is mostly developed around horizontal (ethnic) inequalities and civil wars (Ray and Esteban, 2017). There is no well-defined model linking vertical (class) inequality and social unrest (Esteban and Ray, 2008, 2011). As social unrest and civil wars are different by nature, it is misleading to infer from the literature on civil wars the measure of inequality that would generate social unrest.

In this paper, we study whether inequality generates social unrest at a disaggregated (village) level in India. We start by developing a theoretical framework where social unrest is rationally used by citizens to oppose (unfair) revenue sharing by the elite (e.g. landowners and policy makers) in the village. Our theoretical framework belongs to the class of rational conflict models in which social unrest ensues between the elite and the citizens. The elite decides on how much to concede to the citizens through monetary transfers (wages) and access to a public good (irrigation, pasture). The citizens can resort to unrest, albeit at a cost, to enforce a higher redistribution. This setup gives rise to an equilibrium where unrest increases with a specific measure of inequality: the (log) sum of distances from monthly individual income to

[1]The *Guardian* wrote in 2020: "Most political unrest has one big root cause: soaring inequality" (https://www.theguardian.com/commentisfree/2020/jan/24/most-political-unrest-has-one-big-root-cause-soaring-inequality).

[2]See for example https://www.bloomberg.com/news/articles/2021-04-01/imf-warns-of-social-unrest-trust-erosion-as-inequality-worsens.

the (log) monthly average income in the village for all the individuals whose average income is below the village average. This measure of inequality is close to the well known "mean log deviation" measure of income inequality, used widely in public and development economics (Haughton and Khandker, 2009). Our measure implies that two events can generate unrest: A deterioration of the conditions of the poorest (often referred to as the "opportunity cost" effect, see e.g. Dal Bó and Dal Bó (2011)), or a relative improvement for the richest in the village (a "discontent" effect that increases the incentives to claim for redistribution). While those effects are complementary in the model, we disentangle them in the empirical part.

We test our theoretical predictions using a unique database containing bank account details for individuals across 2,197 Indian villages between January 2015 and May 2017. As a result of India's financial inclusion policies, a formal bank started operating in each of those villages, which greatly improved financial inclusion. Within this context, Somville and Vandewalle (2022) conducted a field experiment that randomly encouraged individuals to open a bank account when the bank had just started operating. They show that the participants who were encouraged to open an account engage in pro-cyclical saving in the account (they deposit more when income spikes). Following these results, we postulate that changes in savings on the account reflect changes in the household's incomes, implying that inequality in savings proxies income inequality. To the best of our knowledge, this is the first paper on inequality and conflict using such a fine-grained database, especially on the time-dimension (the usual measure is expenditure inequality, which is not available on a monthly basis). This feature is particularly relevant in our context as social unrest is typically short-lived and spatially confined. The combination of a very fine temporal and geographical resolution therefore allows us to capture that social unrest events tend to flare up and abate quickly.

Our analysis at the village-month level allows us to highlight how changes in inequality within villages explain the outbreak of social unrest. To overcome a possible reverse causality issue (social unrest might impact the savings capacities), we follow our theoretical setup and implement an instrumental variables strategy that exploits local climate variations. We find that villages with a high level of inequality—measured as the distance between the (log) average and the lowest savings— are more likely to experience social unrest. The effect is sizable: A 10% increase in inequality increases by 6.5% the unconditional probability of social unrest. Furthermore we highlight that the effect on social unrest is driven by an improvement in the surplus of the richest, which is consistent with the discontent motive.

Our identification assumption is grounded on the fact that (positive) water shocks disproportionately benefit landlords and the richest members in a village. We therefore start by providing evidence of a clear differential impact of water shocks on the savings distribution within villages: A larger water availability allows the richest to save more on their bank account.[3] Interestingly, by splitting our measure of water availability between water abundance and drought, we observe that drought is not significantly correlated with bank account savings. The lack of impact of bad water shocks is not surprising in India as (i) Fetzer (2020) demonstrates that the roll-out of the MGNREGA program made the relationship between rainfall and agricultural wages substantially weaker[4] and (ii) Jayachandran (2006) shows that agricultural wages are less sensitive to agricultural productivity shocks in areas where the banking sector is more devel-

[3]The relationship between water availability and agricultural wages and output in Indian villages is well-established in the literature (see e.g., Fetzer, 2020; Jayachandran, 2006; Vanden Eynde, 2018). Furthermore, it has been shown that the variance in landownership is high and that it is the main driver of income inequality in India's agricultural sector (Chakravorty et al., 2019).

[4]MGNREGA is the world's largest employment program, guaranteeing 100 days of minimum-wage employment per year for each rural household.

oped. All the villages in our sample provide MGNREGA and have access to banking through a local banker during our period of analysis. Therefore, consistent with the main theoretical framework, only the rich benefit importantly from water abundance.

Contribution and related literature. First, we contribute to the literature on vertical versus horizontal conflict. Theoretically, the literature on "class conflict" highlights the impact of income inequality on large civil conflict and democratization processes (Acemoglu and Robinson, 2000, 2006; Boix, 2015; Moore, 2016; Scheidel, 2017). Acemoglu and Robinson (2000, 2006) argue that economic elite agree to democratize in societies with medium levels of income inequality in order to prevent rebellion by the poor, thus ensuring political stability. This fundamental model of democratization is ill-adapted when studying social unrest as (i) violence is unrealized at equilibrium and (ii) the model is targeted toward explaining large institutional change. Recently, the literature has highlighted the role of ethnic divisions to explain civil conflicts (Esteban et al., 2012; Esteban and Ray, 2008; Huber and Mayoral, 2019; Laurent-Lucchetti et al., 2019; Montalvo and Reynal-Querol, 2005; Ray and Esteban, 2017). Ethnic divisions matter for large "political" conflicts where the objective is to take over the state apparatus and shaping social and cultural policies (education, religious institutions, infrastructure, etc.). A crucial dimension in this literature is the dissociation of *within* and *between* groups inequality. This dissociation matters as fighting takes both resources and labor (Esteban and Ray, 2011). The argument is that you need both rich people (to finance capital) and poor people (labor and fighters) within a group to fight a war: more inequality within a group might explain conflict. However, similarity between groups increases the chance of war as well, as groups fight for the same resources (competition effect). In this paper, we study the "class conflict" narrative, focusing on social unrest instead of large political conflicts such as civil wars. The crucial difference is that social unrest requires few resources and limited organization: low income is not an impediment to this type of violence (contrary to civil wars). Within-group dynamics between the rich and the poor are therefore likely to be of lesser importance, while the effect of inequality on mobilization still plays a role. We provide a theoretical framework adapted to the problem at hand in order to shape and guide the empirical estimation.

Second, our study contributes to the literature that links income shocks and conflicts. This string of research provides evidence of two opposing reactions to the presence of natural resources: It increases violence through a predatory effect or it decreases violence through an opportunity cost effect (Blair et al., 2021; Dal Bó and Dal Bó, 2011). A string of papers highlight that conflicts around resources are triggered by price shocks (Berman and Couttenier, 2015; Berman et al., 2017; Dube and Vargas, 2013). Finally, weather shocks increase conflict in agricultural areas (Almer et al., 2017; Couttenier and Soubeyran, 2014; Harari and Ferrara, 2018) and social insurance negates this effect (Fetzer, 2020). We complement this literature by showing that an increase in the relative income of the richest members of a society might trigger social unrest for redistributive motives (even if the income of the poorest members does not decrease). An additional contribution to this literature lies in the precise measurement of the effect of water availability on household income. In our context, negative water shocks have no effect on the savings of the poorest while positive water shocks disproportionately benefit the richer households (e.g., the landowners). The usual presumption that water shocks impact social outcomes through a decrease in the income of the poorest members in a society is therefore not validated in this case.

The last strand of literature we contribute to is a set of papers on conflict and inequality in India. Mitra and Ray (2014) analyse Hindu-Muslim conflicts, showing that an increase in Muslim incomes (measured by per capita expenditures) raises violence against them. Another strand of papers focus on Naxalite conflicts, highlighting the role of water availability, forest cover and social insurance in explaining this form of redistributive conflicts (Fetzer, 2020; Ghatak and Vanden Eynde, 2017; Vanden Eynde, 2018).

2 Theoretical Framework

2.1 The setting: Basic structure and timing

We consider a simple static economy (a village) that consists of two groups, the elite (E) and citizens (C). We consider E as a single and homogeneous entity while C is a set of $N = \{1...n\}$ citizens. The elite is the unique owner of productive land and it hires citizens to produce the aggregate village output, generating a total revenue Π. For simplicity, in order to focus on the distributional decisions, we assume that the total revenue of the village $\Pi = \overline{Q} \times P$ is exogenously determined. In particular, we assume that it is composed of the village output $\overline{Q} = f(\overline{K}, \overline{L}, \tilde{\nu})$, generated by the combination of a fixed land size $\overline{K}$, an inelastic labor supply $\overline{L}$ and a (random) quantity of water supply $\tilde{\nu}$, as well as a price P for the output (determined on a competitive market). We also assume that $\frac{\partial \overline{Q}}{\partial \tilde{\nu}} \geq 0$, a higher quantity of water increases the overall output and therefore the total revenue Π. The elite has the power to allocate the village revenue in each period (as it controls the land) and citizens can oppose the revenue sharing scheme through social unrest (riots and protest). The interaction between the elite and the citizens is captured through a game in three stages:

Stage 1: After observing Π, the elite decides on revenue sharing by fixing the individual wages $w = \{w_i\}_{i \in N}$ and the total investment in a local public good $\mathcal{B}$ (use of common land, irrigation scheme, road, education program, etc.).

Stage 2: After observing the proposed revenue sharing scheme, the citizens decide on their social unrest effort (to contest the allocation, at a cost).

Stage 3: Payoffs realize and consumption takes place.

The individual payoff of citizen i is equal to

$$\mathcal{U}_i = \log(w_i) + b_i + \mathbb{I}_{\texttt{Unrest}} \times \theta - c(r_i) \tag{1}$$

where w_i denotes the individual wage; b_i corresponds to the benefits from consuming a share of a common good financed by the elite; and $c(r_i)$ represents the cost associated to unrest efforts r_i (opportunity cost of time, risk of suffering retaliation, etc.). We assume $c(r_i) = \frac{r_i^2}{2\phi}$. Finally, θ is the common psychological benefit of participating to a successful protest against the elite ($\mathbb{I}_{\texttt{Unrest}} = 1$). We interpret the parameter θ as reflecting the social discontent of citizens: If the social discontent against the elite is high, the citizens enjoy a greater benefit from participating in a successful social unrest.[5]

The payoff of the elite group is equal to

$$\mathcal{U}_E = \Pi - \sum_N w_i - \mathcal{B} \tag{2}$$

[5] One can easily extend the model by adding the wage w_i to $c(r_i)$ to explicit the "opportunity cost" channel of rioting. However, it would unnecessarily complexify the model. The analysis will make clear that wages already impact social unrest at the equilibrium (lower wages will increase unrest).

where $\mathcal{B}$ is the total cost of investment in local public goods. For simplicity, we assume that the elite has all bargaining power and set w_i equal to $\overline{w}_i$, the exogenous outside option of the citizens (e.g. the MGNREGA wage). Therefore, the elite only optimizes the amount $\mathcal{B}$ to allocate to local public goods.[6] In section 4.2, we show that the citizens' wages (the low wages in the village) are insensitive to water availability. The theoretical assumption that redistribution happens through alternative means (and not through wages) is, therefore, consistent with our empirical findings.

The citizens may resort to social unrest to contest the revenue sharing rule. Specifically, each citizen can provide a (costly) amount of social unrest effort r_i, with $R = \sum_N r_i$. The social unrest is "successful" if $R \geq \alpha$, where α is uniformly distributed over the interval $[0, \frac{1}{\gamma}]$.[7] In this situation, we assume that the social pressure is so strong that the elite is unable to enforce its announced revenue sharing rule and redistributes the full aggregate revenues for this period: $\sum_N w_i + \mathcal{B} = \Pi$.[8] If $R < \alpha$ the event of unrest fails. This assumption reflects the fact that social unrest is prone to coordination issues, i.e. a protest fails if people are too few and do not protest at the same place, or if the press does not cover the event (implying there is no pressure on the local elite). We assume for simplicity that the elite sets $\mathcal{B} = 0$ when the riot fails (i.e. to punish the citizens).

We further assume that the social discontent $\theta \in \{0, +\mathcal{D}\}$ is not perfectly observable by the elite and is only known to the citizens. We denote by $\mu \equiv \mathbb{P}(\theta = +\mathcal{D})$ the belief that the state of discontent is high. This is a crucial assumption as the elite does not know the value of θ when deciding on $\mathcal{B}$. The elite takes its decision based on μ only. This assumption reflects that leaders may have imperfect information about the exact support they have in the population (see e.g., Wintrobe, 1990). They, therefore, have to take a decision under uncertainty regarding the exact level of discontent. In the model this asymmetric information is the root of bargaining failures and the rational for the existence of social unrest.

2.2 Analysis

We focus on the Bayesian Perfect Equilibrium of the game. To solve the model, we go backwards from the unrest decision by the citizens towards the sharing rule decision by the elite.

2.2.1 Decision on social unrest

Each citizen i decides to participate in social unrest or not. If she participates, she selects the r_i maximizing

$$E\{\mathcal{U}_i|\,\Pi, \theta\} = (\log\left(\overline{\mathcal{W}} + \theta\right) \mathcal{P}(r_i + R_{N\setminus i}) + (\log\left(w_i\right)) \left(1 - \mathcal{P}(r_i + R_{N\setminus i})\right) - \frac{r_i^2}{2\phi} \qquad (3)$$

where $\overline{\mathcal{W}} = \frac{\Pi}{N}$, $R_{N\setminus i} = \sum_{N\setminus i} r_i$ and $\mathcal{P}$ denotes the probability of success of the social unrest event. With our functional form assumptions, this leads to

[6]This setup is close to Anderson et al. (2015). They document that the elite in rural Maharashtra are landlords, who provide wages and might offer public goods to villagers, such as access to trading networks. This can easily be extended to other public goods such as education programs, irrigation schemes and the use of common land for pasture.

[7]This simplifying assumption helps with the exposition, as it linearizes the equilibrium and avoids substitutability in unrest efforts.

[8]The assumption of full redistribution simplifies the exposition. The theoretical results are qualitatively robust to other sharing rules (see the "sensitivity analysis" in Section 4.2).

Result 1 (Probability of observing social unrest)**.** *The probability of observing social unrest is equal to*

$$\mathcal{P}\left(\sum_{N} r_i^*\right) = \gamma^2\phi \sum_{\mathcal{R}} \left(\log\left(\overline{\mathcal{W}}\right) - \log\left(w_i\right)\right) + \gamma\phi\Theta \tag{4}$$

where $\Theta = \sum_{\mathcal{R}} \theta$ and $\mathcal{R}$ is the set of citizens rioting (e.g. for which $r_i^* > 0$).[9] This probability is increasing in the sum of the distances $\sum_{\mathcal{R}} \left(\log\left(\overline{\mathcal{W}}\right) - \log\left(w_i\right)\right)$. Interestingly, this quantity is close to the well known "mean log deviation" measure of income inequality (Haughton and Khandker, 2009).

2.2.2 Revenue sharing decision

Perfect information. For expositional purposes, we first consider the case in which the elite knows the value of θ. Recall that the elite only selects the amount $\mathcal{B}$ as wages are exogenous. Denote by b_i^* the lowest possible b_i that avoids unrest for a citizen i. Accounting for our functional form assumptions and summing the conditions for all individuals leads to

$$\mathcal{B}^* = \frac{\phi\gamma^2}{2}\left(\sum_{\mathcal{R}} (\log\left(\overline{\mathcal{W}}\right) - \log\left(w_i\right) + \Theta\right)^2 \tag{5}$$

The elite can decide to fix $\mathcal{B} = \mathcal{B}^*$ and avoid any unrest, or set $\mathcal{B} = 0$ and face riots with certainty.[10] To limit the number of cases, we assume that $\Theta < \Pi - \sum_{\mathcal{R}} w_i$: setting $\mathcal{B} = \mathcal{B}^*$ brings a higher payoff to the elite than setting $\mathcal{B} = 0$. Therefore, with perfect information, the Elite always sets the optimal amount of public good: either $\mathcal{B}^+$ when $\Theta = \sum \mathcal{D}$ or $\mathcal{B}^-$ when $\Theta = 0$, and always avoids social unrest.

Private information. Alternatively, the elite has imperfect information on the value of θ. In this case, the Elite holds a belief μ on the state of discontent being high when deciding on $\mathcal{B}$. The elite now faces the following trade-off: Fixing a high level $\mathcal{B}^+$ of the public good, which avoids unrest in all cases, or setting a low level $\mathcal{B}^-$ of the public good and risking to observe unrest with probability μ. Therefore, we can conclude that

Result 2 (Optimal strategy)**.** *There exists a unique $\overline{\mu}$ such that $\mathcal{B} = \mathcal{B}^-$ if $\mu \leq \overline{\mu}$ and $\mathcal{B} = \mathcal{B}^+$ otherwise.*

Proof. The proof is in Online Appendix A. □

To take stock, social unrest emerges in the model as a result of private information on the exact level of social discontent in the village. If the probability of having social discontent is low (low μ) the elite is more likely to set a low level of contribution $\mathcal{B}^-$ therefore fostering violence. The probability of observing unrest then increases with participation efforts that depend on the sum of distances between the (log) average and the lowest incomes in the village. This quantity increases when citizen's wages decrease (the "opportunity cost channel") or when the high income in the village increases (the "discontent" effect). This is the relation that we explicitly bring to the data.

[9] We assume that optimal rioting efforts are set to 0 if $r_i^* < 0$.

[10] Notice that any $\mathcal{B} < \mathcal{B}^*$ triggers a riot for sure, the best choice is therefore $\mathcal{B} = 0$ in this situation.

A key theoretical result for our empirical analysis is that an increase in the aggregate village output Π, due for example to a positive water shock $\tilde{\nu}$, increases the probability of observing social unrest in villages with a high level of social discontent:

Result 3. *A positive water shock $\tilde{\nu}$ increases the probability of observing unrest:*
$\frac{\partial \mathcal{P}\left(\sum_N r_i^*\right)}{\partial \tilde{\nu}} = \frac{\partial \gamma^2 \phi \sum_{\mathcal{R}}\left(\log\left(\overline{\mathcal{W}}\right)\right)}{\partial \tilde{\nu}} \geq 0.$

as $\overline{\mathcal{W}} = \frac{\Pi}{N}$ and $\frac{\partial \Pi}{\partial \tilde{\nu}} \geq 0$. This result is due to the fact that a positive water shock $\tilde{\nu}$ increases the total revenues of the village, and therefore the elite rent (top incomes in the village), without affecting the low wages (equal to the exogenous outside option of the citizens). It consequently increases our measure of inequality and the overall probability of observing social unrest. It is noteworthy that this result theoretically holds under more complex wage bargaining frameworks as long as the rent of the elite is more responsive to a positive water shock than the citizens' wages, i.e. as long as $\frac{\partial \Pi}{\partial \tilde{\nu}} \geq \frac{\partial w_i}{\partial \tilde{\nu}}\ \forall\, i$.[11]

3 Background and Data

3.1 Data

Inequality. Our measure of inequality is based on bank account balances, obtained within the context of India's financial inclusion programs. In 2004, the *Business Correspondents Model* was created allowing banks to appoint Business Correspondents as intermediaries to provide financial and banking services on their behalf (RBI, 2006). The Business Correspondents' responsibilities are selecting one person per village to become the local banker, and providing training, equipment and assistance where needed. They also provide a customer service for the clients.[12] We collaborate with Basix Sub-K, a financial inclusion company which has been appointed by seven different banks. The allocation of villages to Basix Sub-K is at the discretion of banks, who in turn received priority lists from the Government. We obtained access to the universe of bank transactions (deposits, withdrawals and transfers) that were facilitated by the banking agents in all the villages in which Basix Sub-K operates (2,574 in total) from November 2010 until May 2017.[13] Figure B1 in Online Appendix B displays the location of the villages. Thanks to the transaction details, we create a panel of balances for 774,341 accounts, starting from the date the account was opened. The richness of the data allows the calculation of our theoretical grounded measure of inequality on a monthly basis: $\sum(log(\overline{\mathcal{W}}) - log(w_i))$. This measure is the sum of distances from the monthly average balance on the account to the monthly average balance in the village for all the individuals whose average account balance is below the village average. For changes in inequality in savings to reflect changes in the household's revenues resulting from economic shocks, income and savings should be positively correlated. Within the same context, Somville and Vandewalle (2022) show people indeed engage

[11] This is the case as long as the bargaining weight of the elite is stronger than the bargaining weight of the citizens. As we will show in section 4.2, this assumption is consistent with our empirical framework: High wages correlate positively with water availability while low wages do not react to these productivity shocks.

[12] The impact of these policies are visible in the statistics from the Global Findex Database. While bank account penetration stood at 35 percent only a decade ago (Demirguc-Kunt and Klapper, 2012), the share of banked adults increased to 53 percent by 2014 (Demirgüç-Kunt et al., 2015) and 80 percent by 2017 (Demirgüç-Kunt et al., 2018). The policies also boosted the inclusion of disadvantaged groups. Between 2014 and 2017, the gap in account ownership between the richest 60 percent and poorest 40 percent of the households narrowed from 15 to 5 percentage points (Demirgüç-Kunt et al., 2018). More recent numbers are provided through the PMJDY website: http://pmjdy.gov.in.

[13] Somville and Vandewalle (2018) and Mehrotra et al. (2021) use this administrative data in their analysis.

in pro-cyclical saving in the account.

Data on social unrest. We use conflict event data from the *Armed Conflict Location and Event Dataset* (ACLED) which contains information on conflict events in India from 2015 until 2018. The data contains information about the date, the location (GPS coordinates) and nature of the events.[14] Prompted by our theory, we gather the information on riots and protests to measure social unrest events. For all villages for which we have information on bank deposits, we compute the monthly number of social unrest events that occurred within 15, 20, 30 or 40 kilometer from the village centroid.

Data on climate. We use monthly variation in water availability at the village level measured by the *Standardized Precipitation Evapotranspiration Index* (SPEI) (Beguería et al., 2014). The SPEI takes into account both precipitation and potential evapotranspiration and therefore captures the impact of increased temperatures on water demand. Our main measure is a continuous variable that takes higher values if drought is more severe. We also test the robustness of our results to different measures of drought.[15]

Other data. Finally, we make use of time-invariant village characteristics: We obtain the population density and the percentage of a village covered by agricultural areas using the Census of India 2011.

Sample selection and summary statistics We collapse the information at the *village-month* level from January 2015 until May 2017, the period covered by all our datasets.[16] They are located in 210 districts in 22 states. On average, 14% of our observations have experienced at least one riot or protest within a radius of 20km, representing 57% of the villages. The standard deviations are relatively large, pointing at substantial variation in both inequality and the measure of weather shocks across and within villages (Table C1 in Online Appendix C).

4 The effect of inequality on social unrest

4.1 Identification Strategy

To quantify the contribution of inequality to social unrest, we rely on equation 4 and estimate a specification of the following form:

$$P(\texttt{unrest})_{v,y,m} = \alpha_1 Inequality_{v,y,m} + FE_v + FE_{y,m} + \epsilon_{v,y,m} \tag{6}$$

where $P(\texttt{unrest})_{v,y,m}$ is a dummy indicating that at least one riot or protest occurred within a 20 km radius of the village centroid v in year y and month m. $Inequality_{v,y,m}$ is our measure of local inequality computed as the sum of the difference between the (log) average savings and the (log) lowest savings in a village v and month y, m. Crucially, the richness of our data allows for the inclusion of a set of fixed effects that account for unobserved heterogeneity at

[14]The data have been widely used in the literature: Berman and Couttenier (2015); Berman et al. (2021); Fourati et al. (2021).

[15]The data have been widely used in the literature: Almer et al. (2017); Harari and Ferrara (2018).

[16]From the 2,574 villages in the original dataset, 332 are missing because of the time period restriction and 45 because of the merging with the other data sources.

the village level (FE_v) and the month level ($FE_{y,m}$). Finally, standard errors are clustered at the village level.

This specification is not immune to identification threats. First, compared to the theoretical equation (4), our baseline specification does not include a time-varying measure of social discontent (θ), as it is not directly observable. The inclusion of the village fixed effects only control for the village-specific time-invariant component of social discontent. It is likely, however, that social discontent has an important time-varying component that we cannot observe and that would bias our coefficient of interest α_1.[17] Second, social unrest may negatively impact savings capabilities, through its effect on working days, the operation of businesses or trade. As pointed out in the model, this is also the case if the elite redistributes in response to unrest. Social unrest may also indirectly affect the measurement of savings in the account, as villagers might not have access to their accounts during episodes of unrest. This may, in particular, hold true for the richest members of the village.

To address the endogeneity issues we adopt an instrumental variable strategy. Following the key theoretical Result 3, we use the feature that an increase in the water availability in the village increases the output and our measure of $Inequality_{v,y,m}$. We therefore make use of time-varying information at the village level on water availability ($\tilde{\nu}$ in the theoretical framework) to instrument our measure of inequality. Empirically, the main assumption is that – conditional on the fixed effects – water availability impacts the likelihood of social unrest only through our measure of inequality. We believe the instrument is relevant in our context. First, the relationship between rainfall realizations and agricultural wages and output is well-established in the literature (see e.g., Fetzer, 2020; Jayachandran, 2006; Vanden Eynde, 2018). Second, the variance in landownership is high and is the main driver of income inequality in India's agricultural sector (Chakravorty et al., 2019).[18] Therefore, we expect a higher water availability to benefit the rich more, mirroring our theoretical framework. We provide evidence for this channel in Section 4.2 by showing the differential impact of water shocks on the income distribution of villages. The exclusion restriction is supported by the use of highly disaggregated data, especially on the time dimension: It is unlikely that alternative channels through which water availability may impact unrest, such as its effect on political participation (Fujiwara et al., 2016), impact local unrest in the exact same month of its occurrence. Therefore, our identification assumption relies on the likelihood that incomes are more reactive to water shocks (e.g. within a month) than other potential alternative channels.[19] We further alleviate this concern in our sensitivity analysis by showing that our main results are robust to the inclusion of demanding village $\times$ year fixed-effects that absorb a substantial part of these longer term alternative channels.

4.2 Main results

Preliminary evidence. Prompted by the endogeneity issues highlighted in the previous section, we show how water availability shapes inequality. In Table 1, we show the first stage regression, highlighting that water availability is positively correlated to our measure of inequality (column

[17] This problem also applies to the belief of the elite on social discontent (μ in the model). This belief is certainly time-varying and unobserved in our main specification.

[18] Focusing on the 2003-2013 decade, Chakravorty et al. (2019) estimate a Gini Coefficient of around 0.6 for income. About half of the income inequality can be explained by variance in income from cultivation, which is in turn primarily dependent on variance in landownership.

[19] It is noteworthy that a potential direct effect of rainfall on political gatherings, such as in Madestam et al. (2013), would work in the other direction than the one we expect. Indeed, abundant rainfall is purported to limit unrest by rendering social gathering more difficult. If this is the case, it plays against the main effect in our IV strategy.

1). To investigate which moments of the savings distribution are the most affected, we define nine deciles (columns 2 to 10). To ease the comparison, we standardized the savings in each decile. We show that water availability correlate only with savings in the last two deciles, for the richest members of villages (columns 9 and 10).

Table 1: Water availability and inequality

	Inequality	Decile 1	Decile 2	Decile 3	Decile 4
	(1)	(2)	(3)	(4)	(5)
Water availability	9.06***	-0.54	-0.24	-0.01	0.22
	(1.28)	(0.33)	(0.33)	(0.35)	(0.41)
Sample mean	442.08	27.41	41.28	55.53	75.11
Observations	55426	55426	55426	55426	55426
	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9
	(6)	(7)	(8)	(9)	(10)
Water availability	-0.27	1.06	0.99	9.65**	30.26***
	(0.84)	(1.30)	(3.73)	(4.28)	(6.05)
Sample mean	108.38	156.31	241.50	394.49	850.19
Observations	55426	55426	55426	55426	55426
Village fe	✓	✓	✓	✓	✓
Year-month fe	✓	✓	✓	✓	✓

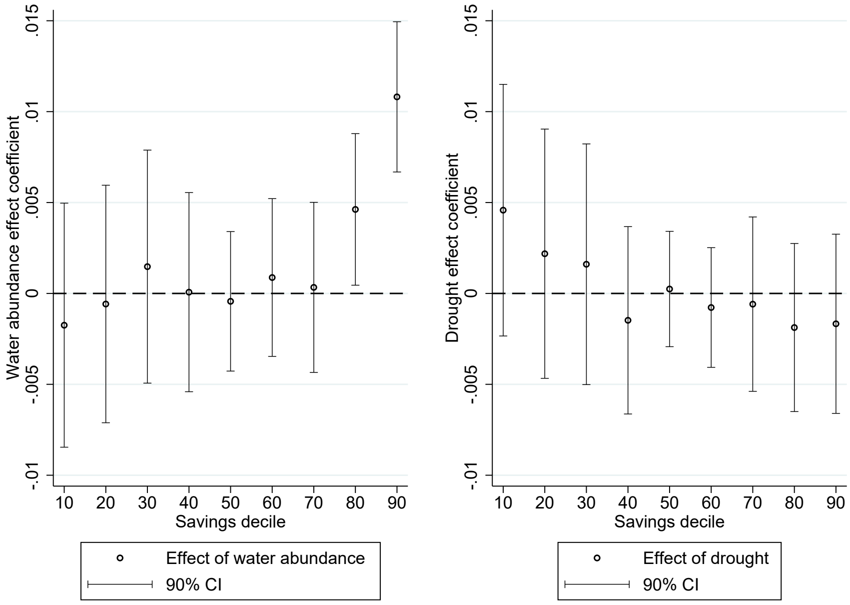
Column (1) presents the correlation between water availability and our measure of inequality, and the columns (2-10) between water availability and the nine (standardized) savings deciles. Standard errors are clustered at the village level and statistical significance is indicated by $* \, p < 0.1$, $** \, p < 0.05$, $*** \, p < 0.01$.

Figure 1 provides a graphical representation of the results when we differentiate between water abundance (spei > 0) and drought (spei < 0). A larger water availability does not impact bank account holdings of the poorest, but allows the richest to have higher savings on their account. Drought is not significantly correlated to bank account savings for any savings decile.

Beside grounding our instrumental strategy, this result constitutes an important contribution to the literature. While a large number of papers have shown the link between adverse rainfall realizations and lower agricultural wages in India, there is - to the best of our knowledge - no evidence on how it impacts the savings buffer of poor and rich households simultaneously. The lack of impact in the lower deciles is not surprising though. First, Fetzer (2020) demonstrates that the roll-out of the MGNREGA program made the relationship between rainfall and agricultural wages substantially weaker. Second, Jayachandran (2006) shows that agricultural wages are less sensitive to agricultural productivity shocks if the banking sector is more developed. All the villages in our sample provide MGNREGA and have access to banking through the BCSA model. Consequently, only the rich benefit importantly from water abundance, which – as pointed out in Section 4.1 – is likely due to the substantial variance in landownership.

Second stage. In Table 2, we present how inequality affects the likelihood of observing unrest. We start by presenting the results of estimating the naive equation 6 (column 1). Our measure of inequality is not correlated with the occurrence of social unrest. As discussed in Section 4.1, this specification is prone to endogeneity issues and we, therefore, present the results of the instrumental variable strategy. The reduced form points to a positive effect of water availability

Figure 1: Water abundance and drought



on the occurrence of social unrest (column 2). The second stage, which is presented in column (3), confirms the prediction by the model: Our measure of inequality has a substantial impact on the likelihood that an event of social unrest occurs in the vicinity of the village in the same month. A 10% increase of inequality increases by 6.5% the unconditional probability of unrest.

Table 2: Baseline estimates

Dep. variable		$P(\texttt{unrest})$	
Model	OLS (1)	Reduced form (2)	IV (3)
Inequality	-0.000 (0.000)		0.001*** (0.000)
Water availability		0.008*** (0.002)	
Cragg-Donald F-stat			34.35
Sample mean	0.14	0.14	0.14
Observations	55426	55426	55426
Village fe	✓	✓	✓
Year-month fe	✓	✓	✓

Column (1) presents the impact on the likelihood an event of social unrest occurs within 20km from the village centroid using an OLS regression. The next columns take into account the need for an instrumental variable: Column (2) presents the reduced form, and column (3) the IV regression. Standard errors are clustered at the village level and statistical significance is indicated by $* \, p < 0.1$, $** \, p < 0.05$, $* * * \, p < 0.01$.

Sensitivity analysis and heterogeneous effect. In Online Appendix D, we present a number of robustness checks. In summary, our findings are robust to i) the inclusion of village $\times$ year fixed effects; ii) changes in the measure of unrest (intensity, winsorizing, different radius); iii) differentiating between drought and water abundance; iv) rescaling inequality; and v) the exclusion of all districts where Naxalite rebels are active.

Finally, we examine the modulating effect of village characteristics on social unrest in Online Appendix E. The positive effect of inequality is mainly driven by villages with a population density larger than the median one. This result is expected, as a higher density translates in a higher number of people rioting in the model (everything else equal) and thus a higher probability of observing an event of social unrest. Furthermore, consistent with our theoretical framework, our effect is larger for villages with more than 60% of the area covered by agricultural activities.

5 Conclusion

We show that income inequalities trigger social unrest in villages in India. To the best of our knowledge, our study is the first to establish a strong empirical link between vertical income inequality and small scale unrest. Taking stock of the fact that the literature on (vertical) income inequality and unrest lacks a well developed theoretical framework, we develop a model where social unrest is rationally used by citizens to oppose (unfair) revenue sharing by the elite in the village. Our main prediction is that the probability of observing social unrest in a village increases with the sum of distances between the (log) average and lowest incomes in the village. We bring this measure to the data by employing unique micro-data on bank account savings covering 2,197 Indian villages. Using our theoretical framework, we propose an instrumental variable strategy grounded on the idea that positive water shocks disproportionately benefit the highest incomes in a village (the landlords). We then document that a 10% increase in our inequality measure increases the unconditional probability of observing social unrest in a given village and month by 6.5%.

References

Acemoglu, D. and Robinson, J. A. (2000). Democratization or repression? *European Economic Review*, 44(4-6):683–693.

Acemoglu, D. and Robinson, J. A. (2006). De facto political power and institutional persistence. *American Economic Review*, 96(2):325–330.

Almer, C., Laurent-Lucchetti, J., and Oechslin, M. (2017). Water scarcity and rioting: Disaggregated evidence from Sub-Saharan Africa. *Journal of Environmental Economics and Management*, 86:193–209.

Anderson, S., Francois, P., and Kotwal, A. (2015). Clientelism in Indian villages. *American Economic Review*, 105(6):1780–1816.

Beguería, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023.

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. (2021). Fertile ground for conflict. *Journal of the European Economic Association*, 19(1):82–127.

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.

Boix, C. (2015). *Political Order and Inequality*. Cambridge University Press.

Chakravorty, S., Chandrasekhar, S., and Naraparaju, K. (2019). Land distribution, income generation and inequality in India's agricultural sector. *Review of Income and Wealth*, 65(S1):182–203.

Couttenier, M. and Soubeyran, R. (2014). Drought and civil war in Sub-Saharan Africa. *The Economic Journal*, 124(575):201–244.

Dal Bó, E. and Dal Bó, P. (2011). Workers, warriors, and criminals: Social conflict in general equilibrium. *Journal of the European Economic Association*, 9(4):646–677.

Demirguc-Kunt, A. and Klapper, L. (2012). Measuring financial inclusion. The global findex database. *World Bank Policy Research Working Paper 6025*.

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., and Hess, J. (2018). The global findex database 2017: Measuring financial inclusion and the fintech revolution. *Washington, DC: World Bank. doi:10.1596/978-1-4648-1259-0.*

Demirgüç-Kunt, A., Klapper, L., Singer, D., and van Oudheusden, P. (2015). The global findex database 2014: Measuring financial inclusion around the world. *The World Bank Policy Research Working Paper No. 7255*.

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

Esteban, J., Mayoral, L., and Ray, D. (2012). Ethnicity and conflict: An empirical study. *American Economic Review*, 102(4):1310–1342.

Esteban, J. and Ray, D. (2008). On the salience of ethnic conflict. *American Economic Review*, 98(5):2185–2202.

Esteban, J. and Ray, D. (2011). Linking conflict to inequality and polarization. *American Economic Review*, 101(4):1345–1374.

Fetzer, T. (2020). Can workfare programs moderate conflict? Evidence from India. *Journal of the European Economic Association*, 18(6):3337–3375.

Fourati, M., Girard, V., and Laurent-Lucchetti, J. (2021). Sexual violence as a weapon of war. *NOVAFRICA Working Paper 2103*.

Fujiwara, T., Meng, K., and Vogl, T. (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics*, 8(4):160–188.

Ghatak, M. and Vanden Eynde, O. (2017). Economic determinants of the maoist conflict in India. *Economic & Political Weekly*, 52(39):69–76.

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

Haughton, J. and Khandker, S. R. (2009). *Handbook on Poverty and Inequality*. World Bank Publications.

Huber, J. D. and Mayoral, L. (2019). Group inequality and the severity of civil conflict. *Journal of Economic Growth*, 24(1):1–41.

IEP (2020). Global peace index report 2020. Technical report.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review*, 109(10):3585–3616.

Laurent-Lucchetti, J., Rohner, D., and Thoenig, M. (2019). Ethnic conflicts and the informational dividend of democracy. *Available at SSRN 3504611*.

Madestam, A., Shoag, D., Veuger, S., and Yanagizawa-Drott, D. (2013). Do political protests matter? Evidence from the tea party movement. *Quarterly Journal of Economics*, 128(4):1633–1685.

Mehrotra, R., Somville, V., and Vandewalle, L. (2021). Increasing trust in the bank to enhance savings: Experimental evidence from India. *Economic Development and Cultural Change*, 69(2):623–644.

Mitra, A. and Ray, D. (2014). Implications of an economic theory of conflict: Hindu-muslim violence in India. *Journal of Political Economy*, 122(4):719–765.

Montalvo, J. G. and Reynal-Querol, M. (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, 95(3):796–816.

Moore, J. (2016). *Injustice: The Social Bases of Obedience and Revolt: The Social Bases of Obedience and Revolt*. Routledge.

Muller, E. N. and Seligson, M. A. (1987). Inequality and insurgency. *American Political Science Review*, 81(2):425–451.

Nafziger, E. W. and Auvinen, J. (2002). Economic development, inequality, war, and state violence. *World Development*, 30(2):153–163.

Ray, D. and Esteban, J. (2017). Conflict and development. *Annual Review of Econonomics*, 9:263–293.

RBI (2006). Financial inclusion by extension of banking services - use of business facilitators and correspondents. *RBI/2005-06/288*.

Scheidel, W. (2017). The great leveler. In *The Great Leveler*. Princeton University Press.

Somville, V. and Vandewalle, L. (2018). Saving by default: Evidence from a field experiment in rural India. *American Economic Journal: Applied Economics*, 10(3):39–66.

Somville, V. and Vandewalle, L. (2022). Access to banking, savings and consumption smoothing in rural India. Working Paper.

Vanden Eynde, O. (2018). Targets of violence: Evidence from India's naxalite conflict. *The Economic Journal*, 128(609):887–916.

Wintrobe, R. (1990). The tinpot and the totalitarian: An economic theory of dictatorship. *The American Political Science Review*, 84(3):849–872.

For Online Publication

Appendix A Theory

Proof of theorem “Optimal Strategy”.

Proof. Notice that if $\mathcal{B}_t = \mathcal{B}_t^+$, the elite payoff becomes:

$$\Pi_t - \mathcal{B}_t^+ - \sum_N w_{i,t} \tag{7}$$

and if $\mathcal{B}_t = \mathcal{B}_t^-$:

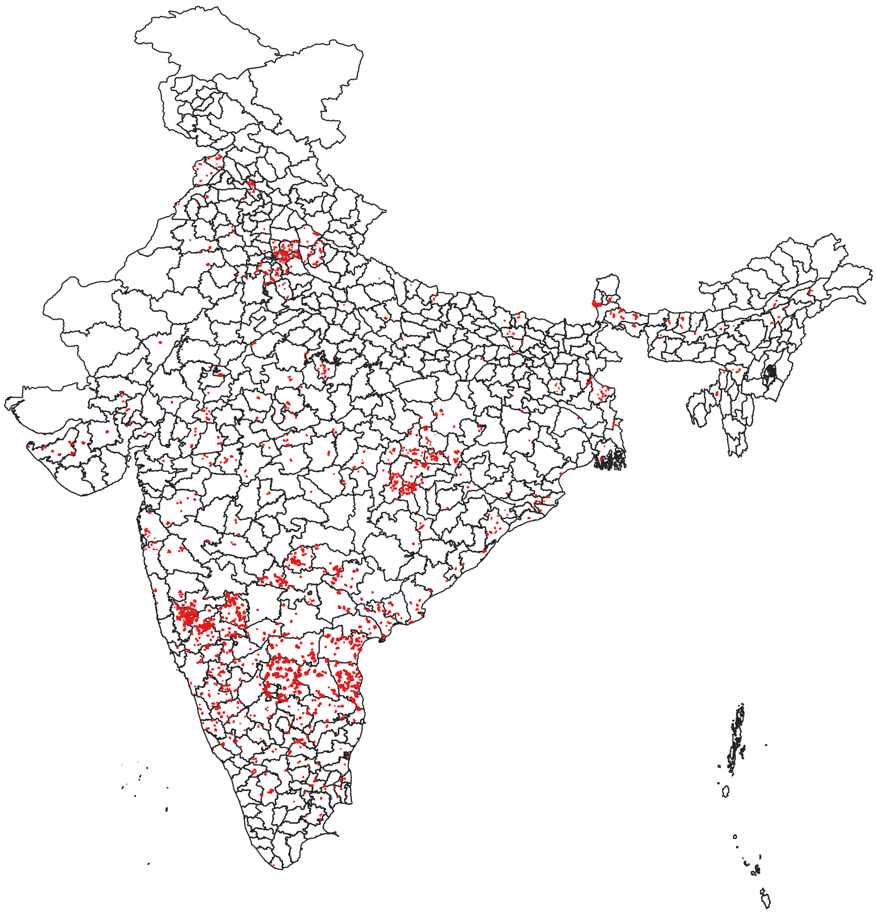
$$\mu \times \left((1-\mathcal{P})(\Pi_t - \sum_N w_{i,t}) \right) + (1-\mu) \times \left(\Pi_t - \mathcal{B}_t^- - \sum_N w_{i,t} \right) \tag{8}$$

therefore when $\mu = 0$, we have that $\Pi - \mathcal{B}^- - \sum_N w_{i,t} > \Pi - \mathcal{B}^+ - \sum_N w_{i,t}$ because $B^- < B^+$ by definition. The elite prefers to set a low level of public good. Furthermore, when $\mu = 1$, $\Pi - \mathcal{B}^+ - \sum_N w_{i,t} > (1-\mathcal{P}) \times (\Pi - \sum_N w_{i,t})$ because $\Theta_t < \Pi_t - \sum_{\mathcal{R}} w_{i,t}$. Finally we have that $\mu \times ((1-\mathcal{P})(\Pi_t - \sum_N w_{i,t})) + (1-\mu) \times \left(\Pi_t - \mathcal{B}_t^- - \sum_N w_{i,t}\right)$ is decreasing in μ. Therefore, by continuity, there exists a unique $\overline{\mu}$ such that $\mathcal{B} = \mathcal{B}^-$ if $\mu \leq \overline{\mu}$ and $\mathcal{B} = \mathcal{B}^+$ otherwise. □

Appendix B   Study area

Figure B1 displays the location of the villages in our sample.

Figure B1: Location of villages in India



Appendix C Summary statistics

Table C1 provides summary statistics for the key variables. Social unrest is frequent: On average, 14% of our observations have experienced at least one riot or protest within a radius of 20km (panel A), representing 57% of the villages (panel B). The standard deviations are relatively large, pointing at substantial variation in both inequality and the measure of weather shocks across and within villages.

Table C1: Summary statistics

	Observations (1)	Mean (2)	Std. Dev. (3)
Panel A: Village-month level			
Occurrence of riots (20km)	55,289	0.14	0.35
Inequality	55,289	158	386
Water availability	55,289	-0.01	0.95
Panel B: Village level			
Occurrence of riots (20km)	2,197	0.57	0.50
Inequality	2,197	138	107
Water availability	2,197	0.04	0.95

Source: Occurrence of riots from *Armed Conflict Location and Event Dataset* (ACLED); Inequality from authors computation using the Basix Sub-K transaction data, and water availability from Beguería et al. (2014).

Appendix D Robustness checks

We first present a number of robustness checks. In Table D1, we investigate whether our estimates are robust to the inclusion of village $\times$ year fixed effects. The results are quantitatively unchanged, which is reassuring given this is a demanding set of fixed effects.

Table D1: Alternative level of fixed effects

	Likelihood of unrest occurring within 20km		
	Reduced form (1)	First stage (2)	IV (3)
Inequality			0.012 (0.019)
Water availability	0.004** (0.002)	0.355 (0.535)	
Cragg-Donald F-stat			0.26
Sample mean	0.14	442.49	0.14
Observations	55370	55370	55370
Village-year fe	✓	✓	✓
Year-month fe	✓	✓	✓

See Table 2 for a description of the different regressions. The regressions include village-year (instead of village) fixed effects, in addition to the year-month fixed effects. Standard errors are clustered at the village level and statistical significance is indicated by $*\, p < 0.1$, $**\, p < 0.05$, $***\, p < 0.01$.

Next, we underline the robustness of our results to changes in the measurement of our key variables of interest. We start by changing the measure of unrest in Table D2. In addition to estimating the impact on whether there was a riot (column 1), we estimate the impact on the number of events (column 2) and on the winsorized number of events at the 99^{th} percentile (column 3).[20] Second, we allow for different ranges of the radius around the village (15, 30 and 40km). Our results remain unchanged. In Table D3, we estimate the sensitivity of our results to alternative measure of water availability at the village level. In panel A, we differentiate between drought and water abundance. In panel B, we follow Jayachandran (2006) and Kaur (2019) and define *water abundance* as the top two deciles, and *drought* as the bottom two deciles in the distribution of water availability per village (per month). Finally, in Table D4, we consider alternative definitions of our measure of inequality. In Panel A, we rescale it by dividing by the total number of account holders. The village fixed effects account for the size of the village, but not for an increasing number of account holders over time. In Panel B, we divide our inequality measure by the number of account holders with a below average balance. This is a more direct measure of the average distance by those who may be inclined to riot. Panel C confirms the results are not driven by outliers, as the estimates are robust to winsorizing the top values at the 99^{th} percentile.

[20] We calculate the 99^{th} percentile conditional on there being at least one event.

Table D2: Alternative measures of unrest

Dep. variable	Dummy (1)	Continuous (2)	Winsorized (3)
Panel A: Likelihood of unrest occurring within 15km			
Inequality	0.000**	0.001**	0.002***
	(0.000)	(0.001)	(0.001)
Cragg-Donald F-stat	34.35	34.35	34.35
Sample mean	0.09	0.25	0.23
Observations	55426	55426	55426
Panel B: Likelihood of unrest occurring within 20km			
Inequality	0.001***	0.004***	0.004***
	(0.000)	(0.001)	(0.001)
Cragg-Donald F-stat	34.35	34.35	34.35
Sample mean	0.14	0.43	0.40
Observations	55426	55426	55426
Panel C: Likelihood of unrest occurring within 30km			
Inequality	0.002***	0.007***	0.008***
	(0.000)	(0.001)	(0.001)
Cragg-Donald F-stat	34.35	34.35	34.35
Sample mean	0.24	0.94	0.90
Observations	55426	55426	55426
Panel D: Likelihood of unrest occurring within 40km			
Inequality	0.002***	0.011***	0.012***
	(0.000)	(0.002)	(0.002)
Cragg-Donald F-stat	34.35	34.35	34.35
Sample mean	0.34	1.81	1.78
Observations	55426	55426	55426
Village fe	✓	✓	✓
Year-month fe	✓	✓	✓

Each panel shows the results considering different ranges of the radius around the village centroid (15, 20, 30 and 40km). In column (1) we show the impact on a dummy indicating at least one riot occurred within the specified distance, in column (2) on the total number of riots and in column (3) on the winsorized number of riots. Standard errors are clustered at the village level and statistical significance is indicated by $*\,p < 0.1$, $**\,p < 0.05$, $***\,p < 0.01$.

Table D3: Alternative measures of *water availability*

Dep. variable	Likelihood of unrest occurring within 20km		
Model	Reduced form (1)	First stage (2)	IV (3)
Panel A: Differentiating between drought and water abundance			
Inequality			0.001***
			(0.000)
Drought	-0.006	4.177	
	(0.004)	(2.839)	
Water abundance	0.011***	19.772***	
	(0.003)	(2.566)	
Cragg-Donald F-stat			28.76
Sample mean	0.14	442.08	0.14
Observations	55426	55426	55426
Panel B: Dummies indicating top two deciles of drought and water abundance			
Inequality			0.000**
			(0.000)
Drought top two deciles	-0.003	2.038	
	(0.003)	(2.573)	
Water abundance top two deciles	0.007**	20.270***	
	(0.003)	(2.507)	
Cragg-Donald F-stat			21.35
Sample mean	0.14	442.08	0.14
Observations	55426	55426	55426
Village fe	✓	✓	✓
Year-month fe	✓	✓	✓

See Table 2 for a description of the different regressions. Each panel shows the results considering different measures of water availability. In panel A, we differentiate between drought and water availability. In panel B, we define *water abundance* as the top two deciles, and *drought* as the bottom two deciles in the distribution of water availability per village (per month). Standard errors are clustered at the village level and statistical significance is indicated by $* \, p < 0.1$, $** \, p < 0.05$, $*** \, p < 0.01$.

Table D4: Rescaling inequality 1

Dep. variable	Likelihood of unrest occurring within 20km		
Model	Reduced form	First stage	IV
	(1)	(2)	(3)
Panel A: Average inequality			
Inequality			1.589***
			(0.565)
Water availability	0.008***	0.005***	
	(0.002)	(0.002)	
Cragg-Donald F-stat			9.69
Sample mean	0.14	0.98	0.14
Observations	55426	55426	55426
Panel B: Inequality per person below average			
Inequality			1.694**
			(0.696)
Water availability	0.008***	0.005***	
	(0.002)	(0.002)	
Cragg-Donald F-stat			6.71
Sample mean	0.14	1.24	0.14
Observations	55426	54694	54694
Panel C: Winsorizing inequality (top 1%)			
Inequality			0.001***
			(0.000)
Water availability	0.008***	7.076***	
	(0.002)	(1.085)	
Cragg-Donald F-stat			27.53
Sample mean	0.14	428.91	0.14
Observations	55426	55426	55426
Village fe	✓	✓	✓
Year-month fe	✓	✓	✓

See Table 2 for a description of the different regressions. Each panel estimates the impact using a different rescaling of our inequality measure. In Panel A, we rescale it by dividing by the total number of account holders. In Panel B, we divide our inequality measure by the number of account holders with a below average balance. In Panel C, we winsorize the top values at the 99^{th} percentile. Standard errors are clustered at the village level and statistical significance is indicated by $* \, p < 0.1$, $** \, p < 0.05$, $*** \, p < 0.01$.

The sensitivity tests in Table D5 are driven by the theoretical framework: The use of the (log) average savings in our main measure of inequality is driven by the simplifying assumption that the elite redistributes fully the revenue in a village after unrest. If less than the full revenue is redistributed (e.g. $s\Pi$ with $0 < s < 1$), we will underestimate the true coefficient α_1 as the proper measure of inequality is $\sum_{\mathcal{R}} \left(\log\left(s\overline{\mathcal{W}_t}\right) - \log\left(w_{i,t}\right)\right)$ with $s\overline{\mathcal{W}_t} < \overline{\mathcal{W}_t}$ and the set of rioters $\mathcal{R}$ diminishes accordingly. We, therefore, investigate in Table D5 if our coefficient of interest changes when we compute our measure of inequality based on $s\Pi$ with decreasing values of s. The magnitude of the coefficient increases as we reduce the share of the revenue that is redistributed in the model, substantiating the idea that the elite redistributes less than the full revenue after unrest (therefore increasing the incentives to riot only for the lowest deciles of incomes in the village).

Table D5: Robustness 5: Rescaling inequality 2

	Likelihood of unrest occuring within 20km			
	First stage (1)	IV (2)	First stage (1)	IV (2)
	Panel A: $s = 0.9$		*Panel B:* $s = 0.75$	
Inequality		0.001*** (0.000)		0.001*** (0.000)
Water availability	8.692*** (1.236)		8.081*** (1.162)	
Cragg-Donald F-stat		33.98		33.52
Sample mean	416.07	0.14	372.90	0.14
Observations	55426	55426	55426	55426
	Panel C: $s = 0.5$		*Panel D:* $s = 0.25$	
Inequality		0.001*** (0.000)		0.002*** (0.000)
Water availability	6.750*** (1.003)		4.758*** (0.754)	
Cragg-Donald F-stat		32.23		30.82
Sample mean	285.76	0.14	167.47	0.14
Observations	55426	55426	55426	55426
	Panel E: $s = 0.1$			
Inequality		0.004*** (0.001)		
Water availability	2.373*** (0.457)			
Cragg-Donald F-stat		24.15		
Sample mean	69.11	0.14		
Observations	55426	55426		
Village fe	✓	✓		
Year-month fe	✓	✓		

Each panel estimates the impact when we compute our measure of inequality taking into account that the elite does not redistribute the full surplus, but only a share s. Standard errors are clustered at the village level and statistical significance is indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$.

Finally, in Table D6, we impose a sample selection: We confirm that our observed link between inequality and unrest is not due to the Maoist insurgency. To do so, we excluded all districts where the Naxalite rebels are active.

Table D6: Other robustness checks

	Likelihood of unrest occurring within 20km		
	Reduced form (1)	First stage (2)	IV (3)
Panel A: Excluding districts with an active Naxalite conflict			
Inequality			0.001*** (0.000)
Water availability	0.008*** (0.002)	7.368*** (1.154)	
Cragg-Donald F-stat			21.61
Sample mean	0.15	448.75	0.15
Observations	51715	51715	51715
Village fe	✓	✓	✓
Year-month fe	✓	✓	✓

See Table 2 for a description of the different regressions. Standard errors are clustered at the village level and statistical significance is indicated by $* \, p < 0.1$, $** \, p < 0.05$, $*** \, p < 0.01$.

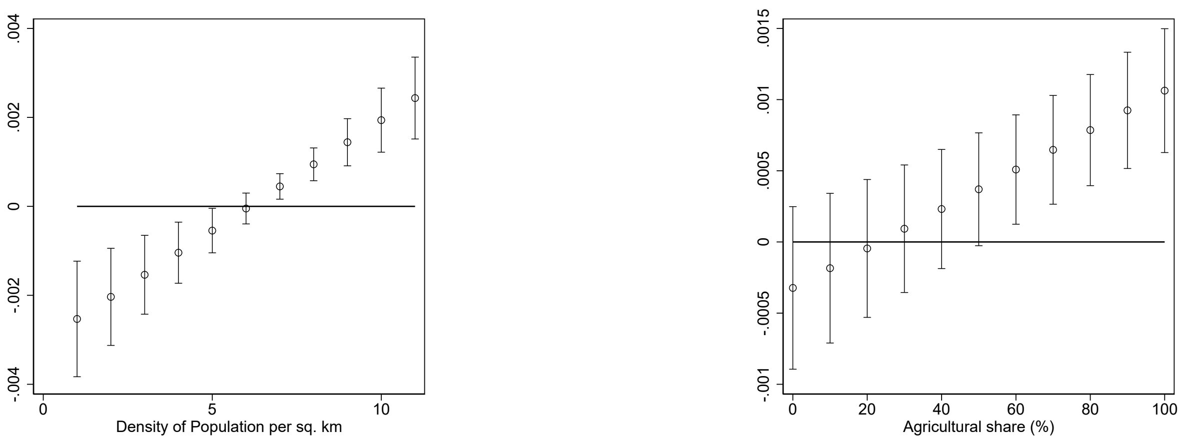
Appendix E Heterogeneous effect

We now examine the modulating effect of village characteristics on social unrest. In Figure E1, we allow the effect of inequality to differ along the line of the following village characteristics: population density and the percentage of a village covered by agricultural areas.

The left panel displays the effect of inequality by bins of population density. The positive effect of inequality is mainly driven by villages with a density larger than the median density. This result is expected as higher density translates in the model in a higher number of people rioting (everything else equal), hence a higher probability of observing an event of social unrest.

Last, our strategy relies on the major assumption that a high share of the economic activity in the villages comes from agriculture. As information on economic sectors is not available at the village level, we approximate the importance of the agricultural sector by considering the percentage of a village covered by agricultural areas. Our effect is much larger for villages with more that 60% of the area covered by agricultural activities.[21]

Figure E1: Heterogeneous effect: Population density and agriculture



[21]More than 90% of villages in our sample has more than 60% of the area covered by agricultural activities