Social Insurance and Conflict: Evidence from India *

Thiemo Fetzer

November 14, 2014

Job Market Paper

Latest version: `http://trfetzer.com/nrega.pdf`

Abstract

Can public interventions persistently reduce conflict? This paper studies whether social insurance is effective in reducing conflict. Adverse income shocks have been empirically and theoretically identified as robust drivers of conflict. An effective social insurance system moderates the impact of adverse shocks on household incomes, and hence, could attenuate the link between these shocks and conflict. This paper shows that a public employment program in India provides social insurance. The program guarantees 100 days of employment at minimum wages providing an alternative source of income following bad harvests. This has an indirect pacifying effect. By moderating the link between productivity shocks and incomes, the program uncouples productivity shocks and conflict.

Keywords: social insurance, civil conflict, India, NREGA, insurgency

JEL Codes: D74, H56, J65, Q34

*London School of Economics and STICERD, Houghton Street, WCA2 2AE London. Support from the Kondard-Adenauer Foundation, the Bagri Fellowship of the Asia Research Centre and the LSE-TISS-TATA grant is gratefully acknowledged. I am especially grateful to Gerard Padro-i-Miquel for his continued support. I also would like to thank Shilpa Aggarwal, George Akerlof, Rabah Arezki, Tim Besley, Erlend Berg, Konrad Buchardi, Jon Colmer, Mirko Draca, Jon de Quidt, Claudia Eger, Christian Fons-Rosen, Maitreesh Ghatak, Benjamin Guin, Clement Imbert, Sam Marden, Hannes Mueller, Torsten Persson, Daniel Sturm, Oliver Vanden Eynde, Kelly Zhang and seminar participants at LSE, CEP, Oxford University and NEUDC for helpful comments.

1 Introduction

Conflict has become an increasingly important topic in recent decades. Millions of lives have been affected by dramatic episodes of social unrest, some of which have turned violent and have led to civil wars and the failure of states. Between 1946 and 2005 civil wars have claimed an estimated 10.1 million lives. Internal conflicts are currently affecting more than one third of developing countries.[1] Billions of dollars are being spent on military interventions in efforts to contain these spreading conflicts. This often takes the form of providing arms and training to the various fighting groups involved. Two open questions arise from all of this. The first is whether such money could have been spent on public interventions to prevent these conflicts from occurring in the first place. The second is how and where these public resources should be directed. This study shows that social insurance policies play a significant role in social stability and the prevention of conflict and social unrest.

Academic literature that identifies the drivers of conflict can help guide policy making. A number of such studies point to two interlinking empirical regularities that stand out. The first is the observation that low incomes provide a breeding ground for civil conflict (Collier and Hoeffler, 1998, 2004; Hegre and Sambanis, 2006); the second is the finding that adverse shocks to incomes cause new conflicts to break out or lead to an intensification of the existing ones (Bazzi and Blattman, 2014; Dube and Vargas, 2013; Besley and Persson, 2008; Miguel et al., 2004; Fearon and Laitin, 2003). This robust empirical relationship between income shocks and conflict provides a blueprint for policy making. Any public intervention that helps households smooth incomes following adverse shocks has the potential to break the link between income shocks and conflict.

The scope for public interventions that aim to protect households from income- and consumption risks is huge in developing countries. The 35 poorest countries in the world with real GDP per capita of less than USD 1000 have experienced 2.8 times more volatile growth in consumption per capita compared to the world's 35 richest countries.[2] Yet, the share of public resources devoted to social insurance programs in developing countries is dismal. Data from the International Labor Organization, for the same set of countries, suggests that in the world's 35 poorest countries only 4% of GDP is spent on social-protection programs, whereas in the world's 35 richest countries 20.5% of GDP, on average, is spent on such programs.[3] The recent World Development Report 2014 shows similar figures. One central policy recommendation of that report

[1] A large literature in economics has tried to assess the true social and economic cost of conflict and the many channels through which it operates, such as by deterring human capital investment (Blattman and Annan, 2010; Leon, 2009; Akresh and Walque, 2008), affecting time preferences (Voors et al., 2012), affecting capital investments (Singh, 2013), diverting foreign direct investment (Abadie and Gardeazabal, 2008) or increasing trade costs (Besley et al., 2014).

[2] Computed as simple average using the World Bank Development Indicators studying growth in consumption per capita between 1995-2011.

[3] Computed as simple average of using data on Total public social expenditure as a percentage of GDP collected by the International Labor Organization for the most recent year available for each country. If one only looks at non-pension spending, the shares are 3 % and 13.7% respectively.

is the development of social safety nets. Given the strong link between adverse income shocks and conflict, an effective social insurance program providing state contingent payouts could have the profound effect of breaking the link between income shocks and conflict. This paper provides evidence that social insurance can achieve just that.

For researchers in this critical field, the challenge is to find a context in which the interplay between social insurance and conflict can be studied. First, it is difficult to find a developing-country context in which such an effective policy or program has been introduced. This is not aided by the fact developing countries spend few public resources on social protection. Even if a country spends significant resources on social protection, it is not clear whether these expenditures truly reflect a social-insurance policy or program in form of a state-contingent payout to individuals who are adversely affected by shocks. Given that developing countries classify a broad set of policies as social protection, the focus needs to be refined to include only policies that have the potential to function as social insurance following this definition of insurance. Last but not least, if such a policy has been identified in a country, it is important to bear in mind that the delivery of social insurance may be difficult in countries that experience conflict.

India best meets the criteria for such a study. It provides a unique testing ground in which to study the relationship between social insurance and conflict. First, throughout its history, the country has suffered from many low-intensity intra-state conflicts. These conflicts have been and continue to be endemic. But their intensity has been relatively low, so that the state still functions on many levels. Second, India has put forth many development schemes. Its most recent efforts began in 2006 when it introduced a public employment program through the National Rural Employment Guarantee Act (NREGA). This program has the potential to function as social insurance. This is because when households demand it, the government provides minimum-wage public employment on local infrastructure projects. Third, NREGA is the biggest public employment scheme in mankind's history. Currently, it reaches up to 47.9 million rural households annually, and generates 210 million person-days of employment. On a typical day, 7.7 million workers are expected to show up to work on one of nearly 294 thousand work sites.[4] Due to its scale, NREGA may have an impact on the dynamics of conflict.

The paper has three main findings. First, I show that before the introduction of NREGA there was a strong relationship between local monsoon shocks and proxies of agricultural income. This in turn creates a strong link between income and conflict. These findings complement and reinforce the existing literature studying the relationship between income and conflict in India.

Secondly, I study the relationship between local Monsoon shocks, income and conflict once NREGA has been introduced. The primary focus of this paper is not on levels of conflict, but rather on the relationship that links adverse productivity shocks with

[4]See http://nrega.nic.in/netnrega/mpr_ht/nregampr.aspx, accessed on 14.06.2014.

conflict. I begin by showing that, since the introduction of NREGA, the relationship between monsoon shocks and agricultural wages has become statistically insignificant. But while local monsoon shocks cease to have an effect on agricultural wages, they continue to strongly predict agricultural output. This finding suggests that agricultural productivity and agricultural wages have decoupled since NREGA was introduced. Hence, NREGA plays a major role limiting the pass through of productivity shocks on to wages, while leaving the agricultural production technology unchanged. I take these results to study how the relationship between monsoon shocks and conflict changes. The key finding is that the relationship between monsoon rainfall shocks and conflict has disappeared since the introduction of NREGA. I show that these results are robust to an array of checks. Further, I obtain similar results studying general violent crime and highlight that the results are most pronounced when studying events of conflict where civilians are the targets of violence. This suggests that NREGA helps bring civilians out of the line of fire.

The third finding shows that NREGA actually is successful in providing insurance.[5] I show that public employment under NREGA supports households and smoothes income following adverse shocks. Participation in the program strongly counters the adverse shocks caused by monsoons along both the extensive margin of overall program participation as well as the intensive margin of number of days worked under the program. In addition, I provide a simple quantification exercise, wherein I suggest that 30% of the district-level-income losses that are attributable to adverse monsoon shocks are compensated by the direct NREGA expenditures that flow into a district. This does not capture the indirect benefits households gain from stabilized agricultural wages.

This paper also makes methodological advances. It is the first to use a novel conflict dataset that covers the whole of South Asia and that has been constructed using scalable Natural Language Processing Tools. The use of natural language processing for economics research has been pioneered by Gentzkow and Shapiro (2010). They use language processing to classify and label text. This paper goes a step further. It does not only label or cluster textual data, but actually extracts information and meaning from text. The idea is to emulate the way humans derive meaning from text, exploiting the grammatical structure of sentences. The hypothetical questions asked of a newspaper clipping about a conflict event are: "Who did what, to whom, when and where?" This semi-automated coding procedure makes the process of coding newspaper clippings covering conflict highly transparent so that it can be used to substitute or, at least as a benchmark for human coding of the conflict data. This approach is particularly important for the economics of conflict, where data limitations and coding routines have been identified as an important factor (Blattman and Miguel, 2010), but there are significant opportunities to use such approaches in other fields of economics research.

There is a vast literature that studies how interventions can affect levels of (agri-

[5]This is in contrast to the puzzling finding of low take-up for weather or rainfall insurance schemes (see e.g. Lilleor and Giné (2005) or Cole et al. (2008))

cultural) incomes. Such interventions can be crudely classified into three categories: (1) physical infrastructure, (2) new production technologies, or (3) politically created institutions.[6] There is a much smaller literature that evaluates the extent to which such interventions actually affect the variability of incomes by insulating against adverse shocks.[7] There is an even smaller literature that studies whether such insulation of incomes weakens the relationship between productivity shocks and conflict that is so widely observed.

In this nascent literature, Sarsons (2011)'s paper falls in the first category studying how infrastructure, in this case dam construction in India, may insulate incomes and through that, affect the relationship between productivity shocks and rioting. Her results suggest that while the construction of dams appears to have moderated wage volatility, it did not moderate Hindu-Muslim riots. The main concern with this finding is that the rioting studied in her paper is mainly an urban phenomena. The ability for physical infrastructure to reduce the volatility of incomes is not necessarily given. Hornbeck and Keskin (2014) find that farmers adjust their production technologies to take advantage of irrigation. They find that this increased production levels but not necessarily to lower output volatility. The insurance that is provided by access to irrigation may induce households to take more risks ex-ante in the crop choice.

The work by Jia (2014) falls in the second category. This study focused on the moderating effect of a drought-resistant sweet potato, as a new technology, on the incidence of riots in ancient China. Jia finds that this type of crop persistently reduced the impact of droughts on rioting. The problem here is that such weather resilient production technologies simply may not exist.

My paper is the first to fall into the third category as it evaluates whether a politically created institution, such as India's National Rural Employment Guarantee Act (henceforth NREGA), achieves the goal of insulating personal incomes from negative shocks and, thus, removes the income dependence of conflict.

This paper also relates to the wider literature on the economics of conflict and labor markets. The theoretical foundation in this field is an opportunity cost argument (see Becker, 1968). A productivity shock reduces the returns to labor and may render joining or supporting an insurgency movement incentive compatible. This eventually translates into increased conflict (see Chassang and Miquel, 2009). Shapiro et al. (2011) study how unemployment levels affect levels of insurgency violence in Afghanistan, Iraq and the Philippines. They find no support for an opportunity-cost channel at work. On the other hand, Iyengar et al. (2011) note that increased public expenditures on construction seems to effect lower levels of violence using labor-intensive technologies. Annan and Blattman (2014) present results from a randomized control trial in Liberia. Their

[6]Duflo and Pande (2007) evaluate the construction of dams and its impact on agricultural production in India. Aggarwal (2014) evaluates the impact of road construction, while Donaldson (2013) studies the impact of railroad construction in colonial India.

[7]Burgess and Donaldson (2010) build on the work of Donaldson (2013) to study how trade integration may have cushioned the effect of adverse productivity shocks on famine mortality.

findings indicate that interventions in the provision of training and capital can greatly increase the opportunity cost of becoming a mercenary and, thus, contribute to a weakening in the relationship between shocks and conflict.[8]

There are fewer studies specific to conflict in India. The main focus of this literature is the Maoist (Naxalite) conflict in the east of India. Gomes (2012) investigates the impact of the Maoist movement and the driving forces behind this conflict. Vanden Eynde (2011) and Gawande et al. (2012) established that the level of Naxalite conflict varies systematically with incomes or proxies thereof, suggesting there is an opportunity cost channel at work. This paper builds on to their work as it presents a study of conflict and crime across the whole of India and addresses how the NREGA workfare scheme has weakened the link between income shocks and conflict.

The remainder of this paper is organized as follows: Section 2 provides some background on the context of conflict in India and the NREGA workfare program. Section 3 presents the data and discusses the novel conflict dataset the author has created for this study. Section 4 presents the empirical strategy. In section 5 there is a discussion on the main results and robustness checks are provided. In section 6, I show that NREGA functions as an insurance program and I quantify to what extent it offsets the risks of conflict. The final section concludes.

2 Context: Conflict and Insurance in India

India's many insurgency movements play havoc throughout the country. In particular, the Maoist insurgency is among the most prolific; it stretches across East India and into several districts in the Northeast, which is also known as the Seven Sister States. Here, various insurgency movements continue to strive for political independence from the Indian government. Movements in the states of Manipur and Assam are the most prolific. It is difficult to distinguish between these conflicts due to their geographic proximity and the existing inter-linkages among the region's movements. The Maoist groups have documented ties with insurgencies in the Northeast and, in particular, with the Manipur-based People's Liberation Army (PLA) and the Assam-based United Liberation Front (ULFA).[9]

The Maoist conflict has been a particular focus of the academic literature.[10] This movement began as a peasant revolt against landlords in West Bengal and their extortive labor practices. In May 1968, the "All India Coordination Committee of Communist Revolutionaries" (AICCCR) was formed. This organization became the root for the armed struggle of several organizations, including the Communist Party of India-

[8]This contrasts with Blattman et al. (2014), who find that a Ugandan employment program, despite large income gains, is not correlated with lower levels of aggression or protests.

[9]See Lok Shaba Unstarred Questions 3138 (13.12.2011) and No 1964 (17.12.2013). Ramana (2007) and `http://goo.gl/bB8wLp`, accessed 20.01.2013.

[10]See Gawande et al. (2012); Gomes (2012); Hoelscher et al. (2012); Vanden Eynde (2011); Khanna and Zimmermann (2013); Morrison (2012); Morgan and Reiter (2013).

Maoist (CPI-M). In its present state the CPI-M is the result of the mergers of various groups that began to take place in the late 1990s and the early 2000s. As an organization, the CPI- M consists of a political wing and an armed wing and is considered a terrorist organization. It is estimated that the military wing the People's Liberation Guerilla Army consists of at least 10,000 combatants.

The Maoist movement is also referred to as the Naxalite movement as it originated in Naxalbari village in West Bengal in the late 1960s. Since then it has spread to less developed areas of the rural areas of southern and eastern India. The Maoists are especially prolific in the states of Chhattisgarh, Jharkhand, Bihar, Orissa, and Andhra Pradesh. They are also present in some states in the Northeast, in particular, in Assam, Arunachal Pradesh, and Tripura. In 2006, around one-third of India's roughly 600 districts were considered to be under the influence of left-wing extremism or subject to violence, thus forming a "red corridor" that now stretches across India (see Figure 1). The aim of the Maoists is to overthrow the existing government and to establish a communist state. The union government under Manmohan Singh has announced that Naxalism poses the largest internal security threat to India.[11]

The Maoist movement is most prolific in India's rural districts. Districts affected by left-wing extremism are marked by a level of underdevelopment that is characterized by lower rates of urbanization, higher degrees of illiteracy, and limited access to infrastructure (such as paved roads, electricity, primary education or health care facilities, see Table A1.) These are all factors that leave these districts wide open to the influence of left-wing extremism. The Maoists are also prolific in the jungle districts of Chhattisgarh, Bihar, and Jharkhand, where they control swaths of land and are said to draw support from the high proportion of the local population's tribal people.

The economic livelihood in the districts affected by left-wing extremism is dominated by subsistence farming, sharecropping, or wage employment in the agricultural sector. The states of Jarkhand and Chhattisgarh host most of India's coal and bauxite reserves and, thereby, attract a lot of investment for resource exploitation. The Naxalites are said to extort taxes from mining companies and to intimidate managers of firms. The main source of income in the forested areas is the production of forestry produce, such as the harvesting of tendu leaves, which are used to produce cigarettes. According to data from the National Sample Survey 2001, 64.9% of Indian households directly rely on agriculture as a primary means of income. In states with a significant Naxalite presence, this share is markedly higher. For example, up to 90% of Chhattisgarh's population is employed in agriculture. For the tribal population, Gawande et al. (2012) highlight the relative importance of income from forestry produce. Any shock to local incomes has dramatic consequences to rural livelihoods. This results in very poor developmental indicators, with tribal households displaying significantly higher levels of food insecurity. Nearly 71.6% of tribal households incur a food deficit for 2 to 3 months each year and 79% of tribal children are reported as being anemic (Radhakrishna and

[11]See `http://www.economist.com/node/15579946`, accessed 25.02.2011.

Ray, 2006).

The Naxalites are highly organized. They have ties with insurgency movements in the Northeast with which they cooperate in the procurement of arms and the training of new recruits.[12] The political wing has a Central Committee that makes key strategic decisions across regions, while the Regional and State Bureaus are responsible for organizing coordinated activities, such as strikes. Local Squad Area Committees have a high degree of autonomy on individual operations. The military wing is called the People's Liberation Guerilla Army (PLGA) and has a similar structure to the political wing. A civilian militia operates at the village level. Its members act as informants and provide direct support and shelter for the armed squads. New recruits are typically sent to training camps before they join an active fighting squad in their home district. At these camps they receive basic military training that lasts between 6 and 12 months and equips them with the knowledge they need for guerilla warfare. This includes how to handle rifles and minor explosives, such as hand grenades, land-mines, and improvised explosive devices (IEDs).

Naxalites attempt to gain legitimacy in these impoverished communities. They do this by actively advocating for local interests. Through this, they foster the popular support that enables them to recruit active fighters. A lot of grievances are aggravated by economic shocks. Some examples of the issues that are aggravated by economic shocks are relationships with moneylenders who forcefully demand repayment; grievances surrounding sharecropping arrangements, which leave farmers with little produce of their own; or the low wages that are being paid. In these environments, Naxalites are said to step in and protect the interests of the local people.[13] This could involve launching "famine raids" (Dash, 2006), calls for bandhs (strikes) to push for higher wages (Ranjan and Prasad, 2012), or targeted violence against civilians that are suspected of having become police informers (Vanden Eynde, 2011).

It is in environments of deprivation and hunger where Naxalites are able to actively recruit new fighters, from the local population, without coercion. Anecdotal evidence on the recruiting of insurgents suggest that monetary incentives do matter. Verma (2011) argues that environments of deprivation present ideal situations for "Maoists to step in, by paying a handsome amount of around Rs. 3000 to the young and promising parents that their kids will have food and money." There are accounts that suggest Naxalites use revenues from extortion and the Narcotics trade to pay these fighters monthly stipends of around Rs 1,500 (see Ramana, 2007). This figure is significant when compared to the average agricultural wages that are earned in India's poorest districts, which range between Rs 50 and 70 per day (see Table 1). Higher levels of recruitment will ultimately lead to more violence once new recruits have been trained and sent back to their home districts.

[12]See Lok Shaba Unstarred Questions 3138 (13.12.2011) and No 1964 (17.12.2013).

[13]See discussion in Pandita (2011) relating to money lenders, Deshpande and Shah (2010) for a discussion about the relationship between droughts and farmer suicides in relation to money lending, Prasad (1987) for a discussion of tenancy arrangements and conflict in Bihar.

While these mechanisms are distinct, they also share the common factor of being facilitated by adverse shocks. This generates the widely observed correlation between violence and the lagged monsoon season rainfall that has been observed in the data. In this paper I show that this relationship between local monsoon shocks and conflict (or general violent crimes) has changed with the introduction of the Mahatma Gandhi National Rural Employment Guarantee Act (NREGA).

The NREG Act was passed in 2005. It established a "right to work" by providing rural households with legal entitlement to 100 days of (minimum) wage employment per household, per fiscal year. The Ministry of Rural Development considers the program to be "the largest and most ambitious social security and public works program in the world" (Ministry of Rural Development, 2012). The goal of NREGA is to develop a strong social safety net for vulnerable groups by providing sources of employment when other employment alternatives are scarce (Ministry of Rural Development, 2008). It is also envisioned as a "growth engine for the sustainable development of an agricultural economy, through the process of providing employment on works that address the causes of chronic poverty, such as drought, deforestation and soil erosion. The Act seeks to strengthen the natural resources base of rural livelihoods and create durable assets which have the potential to generate additional employment in the years to come in rural areas." (see (Ministry of Rural Development, 2008)).

The employment scheme that is defined under the act was rolled out, in three phases, between 2006 and 2008. Nowadays, it covers all Indian districts except a few purely urban centers. In the first phase, from the first quarter of 2006 onwards, the program became active in 200 districts. In 2007, 130 further districts were added. In early 2008, the last phase brought the remaining rural districts into the scheme. The rollout order was far from random. Table 1 provides summary statistics for the districts that fell under the different phases. Districts that received NREGA funding in the earlier rounds were those that had significantly lower agricultural output per capita and lower wages. They were also more likely to be considered under the influence of left-wing extremism and more likely to be experiencing a conflict event. These districts also had minimal access to infrastructure, such as roads, healthcare, and postal services. The endogeneity of the rollout makes it difficult for any empirical study that aims to identify the effect of the program on levels of conflict.

The administration of NREGA is decentralized and aims to empower local governance structures down to the Panchayat level, the lowest level of governance in India. A panchayat typically comprises a few villages or hamlets. In the case where people want to work under the scheme, they approach their local panchayat representative and express their interest in working. The Gram Panchayat will issue the household a job card with which to identify itself. Each household can request one job card. NREGA covers the cost of the household's job card, and all adult members of a household who are willing to work under NREGA are registered on it. The panchayat then has to provide work on a public project, within a two-week period, at the given state-level minimum wage.[14] If the panchayat fails to provide work, a daily unemployment allowance (which is below minimum wage), financed by the state government, is to be paid. The projects that provide workers with employment must be in close proximity to the worker's home (at most, 5 km distance). There is additional remuneration for transportation costs or living expenses, while on the work site. NREGA further requires that 60 percent of a project's budget be allocated to wages. Another requirement is that at least one-third of the workers must be female. Finally, it prohibits the use of machines or contractors.

The design of NREGA includes a major push to promote the further development of a financial sector at local and regional levels. It does so by requiring that all wage payments be made through the banking sector or through bank accounts held at postal offices. NREGA income is to be paid weekly by wire transfer to these local post offices or bank accounts. But, to date, this has not been fully implemented across India. This is a big concern due to the corruption that pervades in this scheme (see Niehaus and Sukhtankar, 2013a,b).

The types of projects are decided at the local level. Districts prepare a shelf of projects, upon which each needs to be agreed with the local panchayats. In particular, the act seeks to empower the panchayats by giving them the right to assign priorities to the types of infrastructure projects that meet local needs and preferences. These projects range from drought-proofing land, to micro-irrigation works, rural sanitation, and rural road construction. As NREGA provides a legal entitlement that becomes available when households demand employment, households can use it as insurance against adverse shocks. This allows NREGA to function as a form of social insurance program that provides protection against idiosyncratic shocks, such as those presented by monsoons.

Low monsoon rainfall is robustly correlated with low agricultural output and wages. In such situations, employment on public works may be an attractive outside option. The design of NREGA has appealing features from a public-economics perspective. NREGA employment requires that households exert efforts to mitigate the effects of factors that impede employment, such as monsoons, by working on public infrastructure projects at minimum wages. The minimum-wage component of the program ensures that households that already have ample economic opportunities have no incentive to work under NREGA even when they could credibly signal that they have been adversely affected by a shock. The self-selection of households ensures that NREGA resources are assigned to households that are most vulnerable and have no access to better-paying jobs (see Besley and Coate, 1992; Nichols and Zeckhauser, 1982).[15] Infras-

[14] Minimum Wage Laws existed before the inception of NREGA. These had limited impact as Minimum Wage Laws were either not enforced (Planning Commission, 2008) or simply did not apply as in the case of self-employment which is the predominant form of agricultural employment.

[15] Such targeting is not achieved by any other government schemes that could be classified as providing social protection; in particular, the Public Distribution System or the system of Minimum Support Prices for agricultural produce. Despite a reform in 1997, turning the Public Distribution System into a Targeted Public Distribution System targeting is still extremely poor (Gadenne, 2014).

tructure construction through public employment makes monitoring of the program easier as output is easily verifiable. This can reduce moral hazard problems that could arise from the decentralized nature of program implementation. In addition to the program's direct employment generation, there are also indirect effects that can contribute to its functioning as insurance for local incomes. These indirect effects may be attributable to the types of infrastructure being constructed under the scheme. Micro-irrigation infrastructure, for example, may persistently moderate the rainfall-income relationship.

The scheme under NREGA is the world's largest known workfare program. During the fiscal year 2009/2010, it generated 2.84 billion person-days of employment for 53 million rural households and, thus, benefitted 291 million individuals. For that fiscal year, an expenditure of INR 37,900 crore, or USD 6.3 billion under the scheme, provided each participating household an average of 54 days of work. Out of this, wages amounted to INR 25,500 crore, or USD 4.2 billion, implying additional labor income of USD 79 per household per fiscal year. This stands as significant in contrast to an agricultural output per capita of INR 13,500 or USD 226.[16]

There is no direct evidence suggesting that NREGA does not work as designed in districts that experience conflict. In fact, it is said that for example in some areas, local Maoist units even put up posters urging villagers to claim their right to employment.[17] Aggregate expenditure data suggest that NREGA expenditures occur in districts that are vulnerable to conflict. The 222 out of a total of 543 districts, in my sample, that experienced some form of conflict over my sample period from 2000 to 2012 accounted for more than 50% of the expenditures under NREGA. The mere size of the scheme in relation to any other development scheme means NREGA could possibly have a profound impact on the relationship between local monsoon shocks and income in the agricultural sector and through this, affect conflict. The next section presents the main data sources, used in this study, that supports the empirical strategy and findings of this paper.

3 Data

In order to provide an overarching picture of how NREGA has affected the dynamics of conflict, relevant data was collected from several sources. Three main data sources are highlighted. First, I rely on a vast collection of newspaper reports covering conflict events across India. This text based data is coded using a novel approach relying on natural language processing algorithms. I complement this dataset with official crime reporting. The second dataset constructed for this study is agricultural income and wage data by exploiting a multitude of different data sources. Lastly, I used some novel

[16] Agriculture accounts for 18% of GDP per capita, but employs 51% of the labor force. See World Development Indicators, `http://data.worldbank.org/indicator/SL.AGR.EMPL.ZS`, accessed 22.08.2014.

[17] Hindustan Times, 22.03.2008.

remote sensed rainfall data.

District Level Conflict data Empirical research on the economics of conflict almost always suffers from severe data limitations. This lies in the nature of the subject under study. Typically, places that exhibit conflict are only weakly institutionalized with little official reporting and scant press and media coverage. Blattman and Miguel (2010)'s review cites that the correlation across different civil-war datasets ranges from 0.42 to 0.96, which may explain why empirical results are often not reproducible when using similar identification strategies, but different datasets or definitions that vary (Ciccone, 2011).

For civil war, differences among datasets can be easily reconciled. However, the conflict literature is increasingly moving toward the study of more micro-datasets at finer spatial and temporal resolutions. Researchers are often left with a set of primary data sources, such as newspaper reports or news feeds from wire services. In order for them to be useful for an econometrician, these reports and news feeds need to be translated into a workable dataset that provides counts of conflict events or incidences at a certain spatial and temporal resolution. As these research efforts are decentralized, this could result in many different datasets being coded from the same raw data sources. The datasets need not coincide because researchers apply different coding practices. And this process does not come without problems.

Datasets produced from the above-mentioned sources are not easily comparable as they are open to the kind of subjectivity bias that makes them difficult to compare across individual coding practices and even more difficult to expand when new primary data sources emerge. This paper uses a novel approach to code the violence data for the whole of India; I obtained stories on conflict and violence from 28,638 newspaper clippings collected by the South Asia Terrorism Portal (SATP). The SATP newspaper clippings represent the most extensive and systematic collection of the raw sources that cover conflict in India.[18]

Although these primary sources have been used by many different authors, when studying conflict in India,[19] this paper is the first to use a novel approach of coding conflict-event data derived from these reports for the whole time period and the whole country. The idea is to use computer algorithms for language processing to emulate the way humans would code conflict data. The core unit of analysis is a sentence in a newspaper clipping. For each sentence, key pieces of information are obtained, namely the subject, verb and object, along with the time and location an event took place.[20]

[18] As the SATP presents only data from English language sources, there may be a systematic selection problem as indicated by Gawande et al. (2012). For the purpose of this paper this is not a concern unless the selection is correlated systematically with rainfall shocks over time.

[19] There is a multitude of research papers that have separately hand coded subsets of the primary SATP newspaper clippings covering various Indian states or various time-periods, see Dasgupta et al. (2014); Gomes (2012); Hoelscher et al. (2012); Gawande et al. (2012); Khanna and Zimmermann (2013); Rana (2013); Shrivastava (2014); Vanden Eynde (2011); Buhaug and Wischnath (2014).

[20] Language processing algorithms, developed for the English language, but increasingly for many other

$$\underbrace{\text{Two unidentified terrorists}}_{\text{subject}}\ \overbrace{\text{massacred}}^{\text{verb}}\ \overbrace{\text{six members of a family}}^{\text{object}}$$

$$\text{and}\ \overbrace{\text{left}}^{\text{verb}}\ \overbrace{\text{a seventh}}^{\text{object}}\,\overbrace{\text{injured}}^{\text{verb}}$$

$$\overbrace{\text{at Mangnar Top, Poonch district, on December 31, 2001.}}^{\text{time \& location}}$$

To illustrate, consider the above example of a sentence. For every verb, the process identifies its underlying subject, object, and surrounding meta-information, such as time and location, which are indicated by prepositions or their syntactic position in a sentence. When this processing step is achieved, the data can be further refined. In the above case, we may want to label the perpetrator ("two unidentified terrorists") of the act of "massacring" to be "terrorists" and the subject ("six members of a family") to be "civilians." This allows a further study of the targets of violence and an analysis of the casualty figures.

This routine is an improvement compared to what the existing literature does. First, this approach allows the study of the myriad "acts" reported. Many hand-coding approaches would restrict the analysis ex-ante to a set of verbs that are indicative of violent activities, such as "to kill." In the automated approach, this can be done ex-post. In this way, one is able to include conflict acts that did not involve casualties, such as attacks on infrastructure, where the words "to kill" would not have appeared. Furthermore, there is no limit to the geographic scope. In order to keep the hand coding manageable, some authors have restricted the analysis ex-ante to cover only certain Indian states or they have searched for district names.[21] Human subjectivity is a third concern this approach addresses. As the routine relies on natural language-processing algorithms, it removes from the coding any subjectivity bias that may emerge. A third advantage is the scalability of the routine. For the purpose of this paper, I construct two main dependent variables. The first is an indicator variable, which is simply a dummy variable equal to one. This is used in case there has been any incident in a district, in a given time period, be it violent or non-violent. I will refer to this dummy variable as "the incidence of violence." The second is a broader measure represented by the number of incidents that have occurred in a given time period. I will refer to this as the violence-intensity variable . The resulting dataset is a balanced, district-level annual panel covering the time period from 2000 to 2012.[22]

The spatial unit used in this paper is an Indian district. I use district definitions from

languages as well, achieve very high accuracy rates in providing a correct syntactic analysis of a sentence, see Fetzer (2013) for a detailed discussion.

[21]This is problematic as a common problem in India is that there is a multitude of spelling variations for similar district names. This could result in significant coding errors or omissions.

[22]The data collection through the SATP began only in mid 2000.

the 2001 census. This is because, since then, many districts have either been carved out of existing ones or renamed. I map these according to the 2001 district boundaries. I study conflict in the whole of India, excluding the Kashmir region. Out of this region, there are 222 districts that have experienced variations in conflict intensity over the sample period. The Ministry of Home Affairs classified 130 of these districts as having been affected by left-wing extremism between 2000 and 2005. Of the remaining 92 districts, 45 are located in the Northeast. The remaining districts are spread across the whole of India. Districts classified as being under left-wing extremist influence account for the bulk (52%) of all conflict events recorded. The states of Assam and Manipur account for 41% of all other conflict events. The situation in these states is either one where insurgency movements have close ties with the Maoists or where there is a significant Maoist presence.

Appendix B.1 provides an example of how the algorithm constructs an incident count that is based on individual newspaper clippings. Appendix B.2 compares the resulting dataset with a geo-referenced version of the Global Terrorism Database. My findings show that the semi-automatically retrieved dataset performs extremely well compared with other violence datasets and even with manually coded data that is drawn from the same newspaper clippings. In addition to this newly created conflict data, I complement this paper by also studying crime in India for the same period.? Particular emphasis is placed on violent crimes and crimes against the public order. These data are collected at the district level by the local police and reported to the National Crime Records Bureau. In order to establish the link between monsoon rainfall and agricultural incomes as a proxy for livelihoods, I collect wage- and agricultural-output data at the district level. The next section describes this data.

Agricultural Production and Wages I constructed two datasets to test whether NREGA had an impact on the relationship between monsoon rainfall and agricultural wages or production. I use agricultural wage data from the Agricultural Wages in India (AWI) series, which has been published by the Indian Ministry of Agriculture since 1951. This publication is unique in that is gives monthly wage rates by district (and sometimes even lists data for multiple locations per district). There is also a separate series that presents wages for several categories of labor, by gender. The quality of the data is very poor, however, as a large number of observations are missing or they report flat wage rates throughout. In order to increase the signal-to-noise ratio, I calculate averages, using the data, in order to generate an annual wage series. In Appendix B.6, I detail some of the issues with this dataset. The resulting dataset is an unbalanced annual panel dataset, at the district level, that covers the time period from 2001 to 2010.

Measurements of agricultural production are more reliable. I use data on the annual district-level production collected and published by the Directorate of Economics and Statistics and the Ministry of Agriculture. This data is reported according to the fiscal year, which begins in April and ends in March of the subsequent calendar year. I match the fiscal year to the calendar year to ensure the largest period of overlap.[23] For every district, I only consider crops that have been consistently planted on at least 1000 hectares for the period that the state reports data. I use state-level harvest prices to construct a district-level measure of agricultural output. The resulting dataset is an unbalanced panel dataset that covers the period from 2000 to 2009. The exogenous variation in this paper comes from a measurement of the intensity of rainfall during the monsoon season. As rainfall reporting from ground measurements is potentially problematic in developing countries, I invoke a novel remote-sensed rainfall dataset.

Rainfall data This paper uses data from the Tropical Rainfall Measuring Mission (TRMM) satellite, which is jointly operated by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace and Exploration Agency (JAXA). The satellite carries a set of five instruments to construct gridded rainfall rates at a very high spatial and temporal resolution. These high resolutions allow it to provide more consistent rainfall estimates than any other available ground-based observations. This dataset is also considered to be the highest quality remote sensed rainfall dataset, with global coverage, currently available (see Li et al., 2012 and Huffman et al., 2007). Its ability to pick up the spatial heterogeneity in precipitation has been highlighted and verified in the Indian context by Rahman and Sengupta (2007). These authors have shown that the dataset from this source outperforms e.g. the Global Precipitation Climatology Centre (GPCC) rain gauge analysis data that has been extensively used in economics research.[24]

The data has the advantage of using a consistent methodology and, most importantly, time invariant sources of input data that are derived from satellite-based instruments. This could be important as, in Appendix B.3, I present some evidence that suggests the number of ground-based measurements that feed into the GPCC could systematically vary with varying levels of violence. The daily rainfall from 2000 to 2012 comes at a fine spatial resolution of a 0.25-by-0.25-degree grid-cell size, which is converted into overall monthly rainfall measured in millimeters. For the identification, I focus on the monsoon season rainfall, which I define according to the principal crops grown as outlined in the state-specific Indian crop calendar.27 The monsoon period varies from state to state as the typical onset dates are early May, for the northeast of India, and may be as late as late June, for central India. For most states, the monsoon period ranges from June to September. Rainfall during the period accounts for at least 70% of India's annual rainfall and most of the precipitation volatility. The last important piece for the empirical analysis consists of NREGA take-up and participation data. As I argue that NREGA provides insurance against income shocks, I study how program participation and expenditures are affected by rainfall shocks.

[23]This data is available on `http://apy.dacnet.nic.in/cps.aspx`, accessed 14.08.2013.

[24]For example by Miguel et al. (2004), Ferrara and Harari (2013) and Kudamatsu et al. (2014). My results are robust to using the GPCC data (Schneider et al. (2011)).

NREGA Participation Data I use NREGA participation data derived from the so-called Monthly Progress Reports (MPR) from before 2011 and from the Management Information System (MIS) from 2011 onwards. The key variables I study are extensive-margin participation as the share of households, in a district, that participates under NREGA in a given fiscal year, the days worked per household, and the total person days generated. In order to be consistent throughout, I match the fiscal year that ranges from April to March to the nearest calendar year. I also obtained data on the number and total cost of ongoing projects, where I specifically classified projects for road construction and land development.[25]

I study three major margins of NREGA take-up. First, I measure the extensive margin as a measure of the share of a district's households that receive NREGA employment in a given year. Second, I measure intensive-margin participation as the log of the number of days worked per household. Last, but not least, I consider the total expenditures per district and financial year. Table 1 presents some summary statistics that suggest NREGA participation is most widespread in districts that received NREGA support in earlier phases. The participation rate in these districts, at around 40%, is almost twice as high as those districts that accessed the program in later periods. Expenditures per capita in districts showed that the districts that received the earliest program spending also received significantly more money. In the first two phases, the spending was around 480 INR per year, per person, compared to only 247 INR per year, per person for the richest districts that received program spending in the last round. The next section presents the empirical strategy.

4 Empirical Strategy

The empirical strategy of this paper consists of linking the three variables monsoon rainfall, income and conflict. I study these relationships for the periods both before and after the introduction of the NREGA workfare program.

First, I analyze the effect of the monsoon season rainfall on agricultural output and wages. I do so by estimating the relationship between agricultural output, wages, and the monsoon season rainfall using an unbalanced panel that covers the time period prior to the introduction of NREGA. This ensures that the estimates are not affected by any impact NREGA might have. I focus on the monsoon season rainfall. The rainfall during this season is most important for India's agricultural productivity. The estimating equation is:

$$log(y_{dprt}) = a_d + b_{prt} + \theta \times log(R_{dprt}) + \epsilon_{dprt} \tag{1}$$

where d stands for district; p stands for the NREGA implementation phase, which ranges from 1-3; r indicates region; and t indicates time. Two sets of fixed effects are

[25]Refer to Appendix B.7 for further discussion of the available NREGA participation data.

included in the regressions. First, there are district fixed effects, ad , which absorb any time-invariant district characteristics that may explain the varying levels of agricultural productivity including soil characteristics, elevation, or terrain ruggedness. The second set of fixed effects are the time-effects bprt. These time fixed effects are region- and NREGA-implementation-phase specific and, thus, remove region-specific time shocks that affect districts that received NREGA support in the first round that differed from that received by districts in rounds two and three.[26] The coefficient of interest is θ. This coefficient measures the elasticity between the monsoon season rainfall and agricultural wages or GDP. For the specifications with agricultural wages, I include a data set that is stated by NREGA-phase-specific linear time trends.

In the second step, I empirically establish the link between the monsoon season rainfall and conflict. I study the two margins, conflict incidence and conflict intensity. Conflict incidence is an indicator of whether there any conflict event has been reported in a district and year. Conflict intensity is the number of all incidences reported in a district and year. The specification that uses conflict incidence is a linear probability model with the estimating specification being

$$A_{dprt} = a_d + b_{prt} + \eta \times log(R_{dpr,t-1}) + \epsilon_{dprt} \tag{2}$$

The fixed effects are as before. η is the coefficient that measures the link between monsoon rainfall and conflict.

The choice of empirical design closely follows the existing literature on conflict in India, which has found a lagged effect of income, or proxies of income, on the intensity of Naxalite conflict (see Vanden Eynde, 2011 or Gawande et al., 2012). The choice of timing can be rationalized by reason that agricultural activity runs in cycles. Periods of peak labor demand are the planting season, at the onset of the monsoon season, when, for example, rice plant seedlings are transplanted to the fields, and the harvest season stretches from October to December (see Figure B2). Incomes are only realized at the end of the year. This holds true particularly for self-employed farmers, which account for 58% of the rural employed across India (see Planning Commission, 2005). A good monsoon creates the chance that a second crop will be grown during the dry season, between November and May, in the subsequent year. All this implies that household income (and expectations) at the beginning of each calendar year strongly depends on the monsoon rains of the previous season.

The second empirical specification uses a conditional fixed-effect Poisson model, as in Silva and Tenreyro (2006), to study the intensity of violence. This type of model accounts for the fact that the number of conflict events is a count variable.[27] The speci-

[26] I define three regions: the North-East, comprised of Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Tripura and Sikkim, the Naxalite Red Corridor, formed by the states Andhra Pradesh, Bihar, Chhattisgarh, Jharkhand, Karnataka, Maharashtra, Orissa, and West Bengal, and a third region comprised of the remaining states consisting of Gujarat, Himachal Pradesh, Haryana, Kerala, Punjab, Rajasthan Tamil Nadu and Uttaranchal.

[27] I use a Pseudo Maximum Likelihood Poisson (PPML) estimator as implemented by Silva and Ten-

fication is as follows:

$$\mathbb{E}(A_{dprt}) = \alpha_d \exp\left(b_{prt} + \eta \times log(R_{dpr,t-1}) + \epsilon_{dprt}\right) \tag{3}$$

The coefficient of interest is again η, which is interpreted as an elasticity. The specification is estimated by using a balanced panel for the whole of India. However, districts that do not experience any variation in the dependent variable do not contribute to the estimation of the coefficients. For that reason, the tables will present a varying number of districts across specifications. Specifications 2 and 3 are reduced forms. It is also possible to perform an analysis of the instrumental variables in order to establish the causal link between the monsoon rainfall and conflict.[28] The use of a lagged rainfall as an instrument for a lagged agricultural output alleviates some direct concerns about the validity of the instrument. However, I mainly focus on the reduced form in this analysis.[29]

The above empirical analysis will be presented in one condensed table that presents the impact of monsoon rainfall on agricultural output per capita, wages, and conflict. The main hypothesis is that before the introduction of NREGA, there was a strong correlation between a lagged monsoon season rainfall and conflict.

The subsequent section presents the empirical specifications used to study how the introduction of NREGA lead to a moderation of this relationship. I do this, using the same empirical setup, by adding an interaction term to the previous specifications, where monsoon rainfall interacts with an indicator $T_{dprt} = 1$ in the case where district d receives NREGA from point t onwards. Note that all districts would eventually receive NREGA.

The identifying assumption for this paper is that the timing of the introduction of NREGA was not correlated with any other omitted variables that could also affect the relationship between monsoon rainfall and conflict. This identifying assumption is valid even if the timing of the introduction of NREGA was endogenous to the levels of violence. The inclusion of NREGA implementation phase specific time effects implies that all identifying variation is coming from districts within the same NREGA implementation phase.

I proceed by studying the moderating effect of NREGA, following the same steps. I first focus on the relationship between monsoon rainfall and agricultural output and wages. The specification that includes agricultural output and wages becomes as follows:

reyro (2006). This estimator overcomes some of the numerical problems in common implementations in statistical packages such as Stata (see Silva, 2011). The PPML estimator does not require the data to have equi-dispersion. It is consistent, so long as the conditional mean is correctly specified. The estimator is even optimal if the conditional variance is *proportional* to the mean, hence over dispersion is not an issue.

[28]See Table A4. The table also presents results for OLS and Negative Binomials as main specification.

[29]An additional problem that arises in particular for the post-NREGA period is the lack of balance in the panel on agricultural output and wages. The data stop in 2009 or 2010 respectively, so there is missing data for many districts that experience conflict both before and after.

$$log(y_{dprt}) = a_d + b_{prt} + \eta \times log(R_{dpr,t}) + \theta \times T_{dprt} \times R_{dpr,t} + \epsilon_{dprt} \quad (4)$$

Note that the simple treatment dummy T_{dprt} is perfectly collinear with the region-by-NREGA phase-time fixed effects. Due to the endogeneity of the roll out, this specification does not attempt to estimate a level effect.[30] The specification asks whether the way rainfall translates into agricultural wages or GDP changes with the introduction of NREGA support. NREGA provides an alternative source of income for households because it provides employment when households demand it. The stabilization of household incomes should also materialize in stabilized agricultural wages as NREGA effectively creates a wage floor and can directly stabilize labor markets. Returning the discussion to the regression, the interest is on the joint significance of the estimated coefficients $\hat{\eta} - \hat{\theta}$.

While the actual employment provision under NREGA makes it reasonable for there to be a direct effect on wage rates, as determined by the agricultural labor market, it is not clear whether there should also be an impact on the relationship between monsoon rain and agricultural output in the short run.[31] In the longer run it is, indeed, possible that NREGA makes agricultural production less sensitive to monsoon season rainfall. This is because the program aims to create infrastructure that benefits the agricultural sector, for example for drought proofing or micro-irrigation. Turning to study the relationship between monsoon rainfall shocks and conflict, the empirical specifications are analogous:

$$A_{dprt} = a_d + b_{prt} + \eta \times log(R_{dpr,t-1}) + \gamma \times T_{dprt} \times log(R_{dpr,t-1}) + \epsilon_{dprt} \quad (5)$$

$$\mathbb{E}(A_{dprt}) = \delta_d \exp\left(b_{prt} + \eta \times log(R_{dpr,t-1}) + \gamma \times T_{dprt} \times log(R_{dpr,t-1}) + \epsilon_{dprt}\right) \quad (6)$$

It is easiest to think of the relationship between the monsoon rainfall and conflict after the introduction of NREGA as an inward rotation: the relationship between monsoon rainfall and conflict becomes less steep. This means that after NREGA is introduced, monsoon shocks of similar magnitude may still translate into conflict, albeit by a smaller amount in comparison to before NREGA is introduced. In the extreme case, the relationship between monsoon rainfall and conflict becomes entirely flat, suggesting monsoon shocks cease to have an effect on conflict.

A common concern with difference-in-difference type estimators is to ensure that common trends hold. I verify this by transforming the treatment variable into a district-

[30] In Appendix section A.2, I explore the level effect of the program as well; however, the identification of a level effect is much more difficult.

[31] NREGA could create income that is used for fertilizer and other agricultural inputs in the short run, which could improve agricultural output. It is questionable however, whether the use of such additional inputs could directly weaken the link between Monsoon rainfall and output.

specific time variable that measures the time to the NREGA introduction. This results in fifteen time-steps.[32] I then estimate a separate coefficient for the monsoon rainfall-conflict relationship for each year. The specifications are as follows:

$$A_{dprt} = a_d + b_{prt} + b_{pct} + \sum_{t=1}^{15} \eta_t \times log(R_{dpr,t-1}) + \epsilon_{dprt})$$

$$\mathbb{E}(A_{dprt}) = \delta_d \exp{(b_{prt} + \sum_{t=1}^{15} \eta_t \times log(R_{dpr,t-1}) + \epsilon_{dprt})}$$

The estimated coefficients η_t can be plotted out along with the upper and lower confidence bounds. The expected coefficient patterns are such that the ?t's are negative for the period before NREGA is introduced and become insignificant for the period after NREGA is introduced.

This paper argues that NREGA breaks the link between monsoon rainfall and conflict because of its moderating impact on household income. When households face adverse conditions, such as those presented by monsoon rainfall, rather than being exposed to income shocks they can earn income through NREGA. This implies that household incomes become less responsive to monsoon rainfall shocks once NREGA is available as NREGA provides a state-contingent payout rather than a lump sum transfer. But whether NREGA functions as insurance is an empirical question. To answer this question I study NREGA utilization along two margins: the overall program expenditures and the total employment it provides in a district over time. The latter is broken up into extensive- and intensive-margin participation.

Extensive-margin participation is measured as the share of households in a district that participate in the program. Since NREGA is available at the per-household level, participation is measured at this level. Intensive margin participation measures the days worked per household. These measures are denoted by P_{dprt}; the estimated specification is as follows:

$$P_{dprt} = \delta_{a_k} + b_{prt} + \phi \times log(R_{dpr,t-1}) + \epsilon_{dprt} \tag{7}$$

The set of fixed effects are similar: b_{prt} are the region and NREGA implementation phase specific time-fixed effects; δ_{a_k} are the district fixed effects that I allow to change for the period beginning from 2011 onwards.[33] The coefficient of interest from these regressions is ϕ. I expect this coefficient to be negative, which indicates that good monsoon realizations are correlated with lower levels of NREGA participation. This

[32]Note that this is longer, even though the panel only ranges from 2000 to 2012. The reason is simple: districts in the first phase have a shorter pre-treatment period, but a longer post-treatment period, while for districts in later phases, this is reversed.

[33] This is because the underlying data-sources for NREGA participation data changes from 2011 onwards, which creates jumps in the NREGA participation data that are specific to each district. Refer to appendix B.7 for more details.

highlights how NREGA take-up is responsive to monsoon shocks. This allows NREGA to function as a stabilizer against adverse shocks to household incomes, thus breaking the direct relationship between local monsoon shocks and income. This is key to explaining why the link between monsoon rainfall, wages, and conflict disappears. The next section presents the results from this analysis and highlights that they are robust to various ways of studying the data.

5 Results

The results are presented in the same sequence as presented in the empirical strategy. I proceed by establishing the relationships between the monsoon season rainfall, agricultural output, wages, and conflict that existed prior to the introduction of NREGA. In the second step, I present the results that pertain to the whole period. This permits the study of how the introduction of NREGA affects the relationship between monsoon rainfall, agricultural output, wages, and conflict.

5.1 The pre-NREGA period: Agriculture, Wages and Conflict

The first section presents the results that pertain to the relationship between monsoon rainfall, agricultural output, wages, and conflict. The regression results are presented in Table 2. Column (1) presents the results that use agricultural output per capita as the dependent variable. The coefficient on contemporaneous monsoon rainfall is interpreted as an elasticity and indicates that a 1% deficiency in monsoon rainfall reduces per capita agricultural output by 0.362%. This indicates a strong dependence of agricultural production on monsoon rainfall.[34] This strong relationship is particularly relevant for self-employed farmers as they are directly hit by adverse shocks. Self-employment in agriculture accounts for at least 58% of rural employment (see Planning Commission, 2005 which establishes a direct link between agricultural production and household incomes). In appendix Table A2 , I show that the production of staple crops is the dominant factor in the relationship between monsoon rainfall and output.

Casual wage employment in the agricultural sector is the second source of income in rural areas. A second step is to analyze whether agricultural productivity shocks, in the form of rainfall shocks, translates into lower wages. This generates a second margin along which productivity shocks can depress household incomes. This phenomenon is studied using the data on agricultural wages presented in the second column of table 2. Again, the coefficient is interpreted as being an elasticity. The effect is small, but significant: a 1% decrease in monsoon rainfall decreases agricultural wages by 0.058%. The effects of monsoon rainfall on the agricultural labor market are highly complex. Yet, this finding relates well with the existing literature (see Jayachandran, 2006). In

[34] This finding is in line with a long standing literature that has studied the relationship between Monsoon rainfall and the welfare of rural households in the Indian context (see e.g. Rosenzweig and Binswanger (1993), Burgess et al. (2011), Cole et al. (2010))

appendix Table A3, I study wages earned in planting as they relate to the harvest season. These two findings are important pieces of information that establish links between rural incomes and rainfall variation.

Next, I address the question of whether a variation in monsoon rainfall explains conflict. The results are presented in columns (3) and (4). Column (3) presents a linear probability model that studies the incidence of conflict in a given year. The coefficient is negative and significant: a good monsoon translates into a lower probability of conflict in a district. A 20% monsoon rainfall deficiency would increase the probability of conflict by 0.7 percentage points. Given that 17.6% of the district years exhibit conflict over the period under study, this shows an increase of 3.9%. Column (4) presents the results from the Poisson regression. Note that the regression is estimated by using the whole balanced panel. However, only 141 districts provide a time variation in the dependent variable and, thus, contribute to the estimation of the coefficients. The coefficient is interpreted as being an elasticity that indicates a 1% reduction in monsoon rainfall translates into a 0.846% increase in conflict. This coefficient compares very well with the estimates presented in previous studies, in particular, Vanden Eynde (2011) and Gawande et al. (2012), who study Maoist conflict. It also establishes a direct link between monsoon season rainfall and conflict.[35]

Highlighting the non-linearities in the relationships between monsoon rainfall, agricultural output, wages, and conflict is also useful. This can be accomplished by estimating these relationships non-parametrically and display the results as in Hsiang et al. (2013). The idea behind using a non-parametric approach is to obtain local estimates of the relationships being studied and to visually display them visually weighted to highlight the degree of certainty.[36] Panel A in Figure 3 presents the relationship between monsoon rainfall and agricultural output per capita. This relationship seems fairly monotone. The second graph depicts the relationship between wages and monsoon rainfall. Again, a fairly monotone relationship emerges, as indicated by the linear fit. The conflict non-parametric fit exhibits some non-linear patterns. The estimated coefficients are positive for monsoon rain deficits, indicating that a deficient monsoon is correlated with a higher level of violence. Large and positive monsoon rain deviations are generally correlated with less conflict. The relationship appears to be bending back up, suggesting that extremely positive rainfalls can also induce conflict. This is

[35]In appendix tables A4 I perform a set or robustness checks highlighting that I obtain similar results using an instrumental variables approach and that the results are robust to the choice of empirical model.

[36]The procedure has two steps. First, the data is demeaned by the fixed effects. This ensures that the scales are identical. A loess regression of the residuals of the weather variable and the dependent variable of interest is then estimated repeatedly using a bootstrapping procedure. The residuals for the horizontal axis, in this case the residuals for (lagged) Monsoon-rainfall, are subdivided into a set of 200 grid points. Each bootstrapped regression is evaluated at the grid-points for the horizontal-axis. This results in a set of fitted values for each grid point along the horizontal axis. In the second step the fitted values are plotted. For each horizontal grid point, a kernel density is estimated. The colouring is related to two things. First, the overall color intensity at each horizontal axis grid point is related to the overall mass of data that accrues there. This color is then streched out vertically in relation to the density of the fitted values. 95% confidence bounds are plotted as dashed lines.

not surprising: the non-parametric for agricultural output seems to be bending down at very high monsoon realizations.[37] In the next section, I study how the introduction of NREGA smoothes out these relationships.

5.2 After NREGA: Moderation of Monsoon and Conflict Relationship

The focus of this paper is how NREGA changes the slope that links monsoon rainfall and conflict. NREGA's ability to provide employment opportunities incomes is the instrumental link through which a moderation of the impact of monsoon rainfall on incomes and indirectly on conflict is achieved. I proceed by presenting the results that show the relationship between monsoon rainfall, agricultural output, wages, and conflict as presented in the previous section. The only addition here is as has already been discussed in the empirical strategyhat the slope that links monsoon rainfall with each of these three variables of agricultural output, wages, and conflict changes with the introduction of NREGA.

The results are presented in a condensed form in Table 3. Column (1) presents the results that use agricultural output per capita as the dependent variable. The focus is on the relationship between the coefficient of monsoon rainfall with the coefficient of the interaction between the NREGA treatment dummy and monsoon rainfall. While the former is positive and significant, the latter is negative and insignificant at conventional levels. This suggests that the relationship between rainfall and agricultural output per capita has not changed, at least up to the year 2009, when agricultural output panel ends. This is not surprising. NREGA aims to produce sustainable local infrastructure, which in the longer run could increase agricultural output and make it more resilient to monsoon rainfall shocks. Nevertheless, these should not have an immediate effect on this deeply structural relationship. It is, thus, not surprising that agricultural output is still very much a function of monsoon rainfall.

Column (2) presents the results for agricultural wages. NREGA is a major intervention in the agricultural labor market; in periods in which agricultural wages would otherwise have been depressed, due to an adverse weather shocks, NREGA provides an outside option that should stabilize agricultural wages and make them less responsive to rainfall shocks. The regression indicates this to be the case. The coefficient on the interaction term is positive and significant. The sum of the two coefficients is insignificant as indicated by the F-test. This implies that agricultural wages cease to be a function of monsoon rainfall after the introduction of NREGA. This indicates that NREGA can exert a general equilibrium or stabilizing effect on agricultural wages for households that do not directly participate in the program. This finding complements the existing empirical research that documents NREGA's effect on increased wage levels (Azam, 2011; Berg et al., 2012; Imbert and Papp, 2014; Zimmermann, 2012).

[37] A general linear fit appears adequate given the data. The slight non-linearity for positive extremes, if anything, implies that I underestimate the steepness as the regression line is pulled back up.

The last two columns show the relationship between conflict and monsoon rainfall. The coefficients in both columns indicate that the monsoon rainfall-conflict relationship in regard to conflict incidence (column (3)) and conflict intensity (column (4)) has been dramatically moderated. In order to statistically assess the degree of moderation, I perform an F-test on the joint significance of the two monsoon rainfall coefficients. The F-test is insignificant with a p-value of 0.267. This suggests that there remains a negative relationship, indicating that low monsoon rainfall translates into conflict. However, the relationship is a lot weaker and statistically insignificant. This suggests that NREGA has completely removed the rainfall dependence of conflict. This relates well with findings in historical China, where the introduction of the drought-resistant sweet potato has moderated the impact of weather shocks on peasant revolts (Jia, 2014).[38]

It is important to highlight that the finding in this study is distinct from the empirical approaches that aim to identify a direct effect of NREGA on conflict levels. Poor districts and districts that have experienced conflict received NREGA support earlier, making districts that received NREGA support later a poor counterfactual. My identification strategy steers clear of this concern. Nevertheless, it becomes instructive to estimate the effect of the introduction of NREGA on levels of conflict.

In Appendix A.2, I use a simple difference-in-difference strategy to estimate the effect levels. The estimated coefficients suggest that the levels of conflict are 30% to 50% lower after the introduction of NREGA. This maps well into the findings of Dasgupta et al. (2014). They use a difference-in-difference estimator to estimate NREGA's impact on conflict levels. The identifying variation this relies on comes from the sequential rollout of NREGA. I show that the estimated effect has two sources: First, NREGA may have a direct effect in terms of the level of conflict as it has led to an increase in wage levels that are independent of weather shocks. In the context of classical opportunity-cost-based models of conflict, this can be seen as pushing out the participation constraint. However, the second margin is an insurance effect that prevents wages from dropping in the case of a bad monsoon.

Khanna and Zimmermann (2013) obtain different results studying conflict levels. They directly address the endogeneity of the rollout and use a fuzzy regression discontinuity design for identification. They reverse-engineer the NREGA roll-out algorithm to identify the districts that were close to the cutoff of being assigned to an earlier or later phase. Districts close to the cutoff serve as counterfactuals. Their results indicate that the introduction of NREGA has led to an increase in the levels of conflict in the short run. They argue that this is due to an increase in civilian collaboration with security forces. This induces more violence in the short run, but a moderation in conflict levels in the longer run. A concern with the research design is that it is driven by relatively few districts that experienced conflict variation and were near the cutoff. While this finding does not square with the insurance effect I document in this paper, it is at

[38]Related is the recent work by Kung and Ma (2014), who show that cultural norms seem to moderate the effect of adverse shocks on revolts in historical China as well.

odds with the (not well identified) level effect I estimate. Further, they find this increase in conflict levels to be immediate. As the NREGA implementation was rather gradual it is not clear whether this is to be expected.

In order to study the dynamics of this effect, an event study analysis can highlight the degree to which the relationship between monsoon rainfall and conflict evolved prior to the introduction of NREGA. The treatment variable is transformed into a district-specific time variable that measures the time back to the introduction of NREGA. Note that this is a longer time period, even though the panel only ranges from 2000 to 2012. The reason for this is that the districts in the first phase have a shorter pre-treatment time period, but a longer post-treatment time period, while this is reversed for districts in the later phases.

The result is a sequence of the estimated coefficients, $\hat{\eta}$, that are best visually presented (see Figure 2). I only plot the coefficients that are estimated by using the districts from all three phases; otherwise, the picture would be distorted due to a compositional effect. The vertical line around zero refers to the point in time when NREGA was introduced. The dashed blue lines indicate the regression coefficients obtained from the baseline specification. The estimated coefficients suggest a consistently negative relationship between monsoon rainfall and conflict prior to the introduction of NREGA. The relationship disappears with the introduction of NREGA. It does so, not instantaneously, but rather gradually, with the coefficient becoming insignificant (only one) roughly two years after NREGA was introduced. The results suggest that common trends hold.

The non-parametric analysis suggested some non-linearities in the relationship, it is important to explore this for the post-NREGA period as well. The results are presented in Panel B of Figure 3. It is best to directly compare Panel A from before the NREGA period with Panel B from after NREGA was introduced. Note that for both graphs the scales are identical, which allows for a direct comparison. The apparent patterns are very similar to those indicated by the linear regressions. The relationship between monsoon rainfall and agricultural output per capita has remained very similar. However, the relationship between agricultural wages and monsoon rainfall has rotated inward: Following the introduction of NREGA, there appears to be no statistically significant relationship that links monsoon rainfall and agricultural wages. This indicates that the stable outside option NREGA provides serves as a cushion for wages that are determined in the labor market. The relationship between monsoon rainfall and conflict follows in the last column. The non-parametrics paint a very suggestive picture, that the relationship becomes flat. But in the period before NREGA, the relationship was weakly U-shaped. This indicates that the relationship between the weather extremes that translate into conflict disappears following the introduction of NREGA. This points to the fact that, since the introduction of NREGA, monsoon rainfall variation has ceased to have an effect on conflict.

Before exploring heterogeneity and the mechanisms through which this moderation in the monsoon conflict relationship was achieved, I perform various robustness checks to highlight that the core result is robust.

5.3 Robustness Checks

I categorize the robustness checks into three sets: The first check involves adding more control variables and time varying fixed effects. The second set uses different data sets or measures of monsoon rainfall. The last set includes some placebo tests.

The first set of robustness checks explore the robustness of the results of the inclusion of different sets of fixed-effects or by adding controls. These are presented in Table 4. The first three columns study conflict incidence, while columns (4) to (7) study conflict intensity. I discuss the corresponding conflict incidence and conflict intensity specifications together. Columns (1) and (4) explore the robustness of the results, to the inclusion of a set of time-invariant district characteristics from the 2001 census that interact with a set of year fixed effects. This is in the spirit of the semi-parametric difference-in-difference analysis as developed in Abadie (2005).[39] The estimated coefficient changes in sign and becomes insignificant for the conflict incidence regressions. However, the results for the conflict-intensity specification is robust. The fact that the linear probability model becomes insignificant is not too concerning as the there is too little variation in the dependent variable. In columns (2) and (6), I introduce time-varying district fixed effects where I estimate a separate set of district fixed effects for the period before and after the introduction of NREGA. The inclusion of such fixed effects would capture any district, specific level effect that the introduction of NREGA may have had on conflict. While it is unlikely that the presence of the scheme triggers a dramatic conflict response, these fixed effects shut down this channel. The insight is that, despite the inclusion of the fixed effects, the estimated coefficients change only slightly for the conflict intensity specification. Again, the interaction for the incidence regressions becomes insignificant, which is not surprising, given that the time varying fixed effects explain most of the variation in the dependent variable. While controlling for the NREGA treatment dummy, columns (3) and (7) introduce state-by-year fixed effects. These fixed effects absorb a lot of the variation in monsoon rainfall. As law and order is in the domain of the Indian states, and not the union government, it is reassuring that the coefficient pattern remains similar. Lastly, column (4) studies only the districts that experienced conflict before the introduction of NREGA. The estimated effect for this subset of districts is very similar to the main specification. This ensures that the effect is not driven by an expansion of the geographic range of conflict that is correlated with the introduction of NREGA and monsoon rainfall.

[39]The district characteristics are: terrain ruggedness, elevation, rural population share, tribal population share, scheduled caste share, illiteracy rates, household size, share of population younger than 6 years, population growth rate from census 1991 to 2001, gender gap, share of villages in district with primary school, share of villages in district with mud road approach, share of households in district that live in permanent housing, share of villages in district with primary health care facilities, share of villages with electric power, share of villages with a bus stop and the share of villages with a postal office.

The second set of robustness checks concerns the measurement of local Monsoon shocks. I present robustness checks using different rainfall data or agricultural productivity proxies. These exercises are presented in Table 5. Again the analysis is presented on conflict incidence and intensity margins.

Columns (1) and (7) present the results from a specification where district-level rainfall is normalized by its standard deviation. This is problematic, as the 14-year time period for which the TRMM data is available may be too short for a stable estimate of the mean volatility to emerge. Nevertheless, the results obtained from using this measure of rainfall are very similar. Columns (2) and (6) present the results from the study of a different rainfall dataset. I present results based on the GPCC data, which is based on ground-level rain-gauge measurements.[40] While I believe that the satellite-based TRMM data is likely better, as rainfall reporting may be endogenous to conflict (see appendix B.3), it is nevertheless reassuring that I obtain very similar results when using a different dataset. As both the GPCC and TRMM data have been processed using climatology algorithms, a general concern is "error propagation" (see Leung et al., 2005; Burnicki et al., 2007). As the raw data is transformed in the analytical process, the mathematical and numerical transformations may propagate small simple measurement errors. This could generate spurious correlations that could affect the results.

A simple way to address this is to instrument one rainfall dataset with the other one. This removes any systematic and non-systematic measurement error and ensures that the results I obtain from the two different datasets are driven by the same underlying variation. The results are presented in columns (3) and (7) and are very similar to the core finding. Monsoon rainfall is only one proxy for local weather shocks. The analysis of the agricultural output relationship suggests that it absorbs a lot of the variation in agricultural productivity. Nevertheless, it may make sense to explore a different measure.

Gawande et al. (2012) propose using a vegetation index that measures the degree of photosynthetic activity as a potential measure of agricultural output. Photosynthetic activity drives plant growth and, thus, agricultural output. It is obviously affected not only by rainfall, but also by other climatic conditions. Columns (4) and (8) present the results from a study of the lagged Normalised Vegetation Index (NDVI) and its interaction with NREGA. The coefficient pattern is very similar to that in the preferred specification. However, the NDVI contains a lot less variation, which makes it more difficult for the coefficients to gain significance. Nevertheless, using this proxy leads to very similar results. These results give me confidence that I am genuinely picking up an impact of monsoon rainfall on conflict.

I now turn to a third set of robustness checks, which explore different placebo tests. I perform two main placebo tests. The first one simply follows the standard approach in

[40]The data has been widely used in the economics literature and in the conflict literature in particular. Some prominent references are Ferrara and Harari (2013), Miguel et al. (2004), and Kudamatsu et al. (2014) and many others.

difference-in-difference type empirical setups: moving the program reform to an earlier date. This allows me to test whether the relationship between monsoon rainfall and conflict had already begun to change before NREGA was introduced. The previous exercise that studied the effect of monsoon rainfall over time already indicates that this is not the case. Moving the treatment to an earlier date should eliminate the estimated NREGA effect.

The second placebo check tests whether rainfall outside the main growing season has some effect on conflict. If monsoon rainfall effects rural livelihoods, then rainfall outside the monsoon season should not in any systematic way interact with NREGA as explaining conflict. The results are presented in Table 6. Columns (1) to (3) and (5) to (7) present the results when moving the NREGA introduction to respectively three, two and one years ahead. Moving the NREGA treatment to an earlier date should wash out the estimated NREGA effect. This is because, in truth, the rainfall-conflict relationship was present for the time period between the placebo NREGA introduction date and its true introduction date. Conversely, the closer the placebo is moved to the true NREGA introduction date, the more likely it is that we should observe an NREGA effect for the placebo. This is the exact pattern that emerges in the data when studying the regression coefficients for conflict intensity in columns (5)-(7).

The conflict incidence regressions have positive and significant coefficients on the Placebo interactions. If I restrict the analysis to only the period prior to the true NREGA introduction, the coefficients become insignificant and flip signs. This suggests that the positive and significant coefficients on the Placebo reform are driven by the period after the actual NREGA was introduced. Columns (4) and (8) present the results when rainfall outside the monsoon season is studied. As expected, the coefficients are insignificant, though they share the same coefficient pattern for conflict intensity. This is likely due to rainfall outside the monsoon season being positively correlated with rainfall during the monsoon. These exercises suggest that the introduction of NREGA fundamentally changes the relationship between monsoon rainfall and conflict.

In the next section I show that this also applies more generally to violent crimes, suggesting that NREGA also has more general indirect effects on the income dependence of crime.

5.4 Effects on Crime

The existing literature highlights that there exists a relationship between weather and crime in the context of India. The common observation is that adverse rainfall shocks drive crimes against vulnerable populations, in particular crimes against populations from scheduled castes and scheduled tribes (see (Sekhri and Storeygard, 2013), (Iyer and Topalova, 2014)), and violence and property crimes in general (see (Blakeslee and Fishman, 2014)). To study the impact of NREGA on the relationship between monsoon rainfall and crime I obtained the same district-level crime data used by these authors for the period 2002 to 2012. I present the results, using the main specification as in the previous section with a lagged monsoon rainfall. I use the log of the number of crimes reported by broad categories as the dependent variable.[41] The results are presented in Table 7. The coefficients suggest a moderation of the crime-rainfall relationship, most prominently for violent crimes and for disruptions of the public order. This crime category includes incidences of rioting and arson. These results map well into my findings on insurgency-related violence. Some of the events captured in the conflict data may also be measured in the violent-crime data. The results are driven by the districts that are under the influence of left-wing extremism and districts that received NREGA support in earlier phases. This maps very well into the results found when using the conflict data.[42]

A key concern with the crime data is that data from cities is highly overrepresented, especially for property crimes. To illustrate: the state of West Bengal has about 91 million inhabitants, and the capital, Calcutta, has roughly 4.6 million inhabitants, accounting for about 5% of the population. In the raw crime data for 2005, Calcutta accounted for 21% of all thefts in West Bengal and about 3% of all murders. This suggests a strong urban bias for crime reporting.[43] This is not, in itself, a problem. But if the relationship between crime and monsoon shocks is different for cities relative to rural areas, this could wash out the variation in the crime data that is attributable to monsoon shocks.[44]

For violent crimes and crimes against the public order, in particular murders and rioting, the reporting bias is weaker as these are highly visible crimes. A non-parametric analysis for the data on violent crime suggests a pattern that is very similar to the one observed for the conflict data: Before the introduction of NREGA support, extremes in monsoon rainfall in the preceding year translate into more violent crimes. This relationship is significantly moderated, following the introduction of NREGA: the U-shape becomes flat (see Figure 6). Estimating the effect of a lagged monsoon rain over time suggests a similar response: negative coefficients for the period before NREGA was introduced and insignificant coefficients following NREGA's introduction (see Figure 7).

Next, I study some heterogeneity in the estimated NREGA effects. In particular, I

[41]Poisson models as used in the rest of the paper yield very similar results and are available upon request. I follow the categorisation of Iyer et al. (2012), Appendix 1. The only modification is that I do not consider only murders, but violent crimes includes the crime categories: murder, attempted murder, kidnapping and hurt.

[42]My results differ from those presented in Iyer and Topalova (2014) who do not find any systematic moderating effect of NREGA on the crime- and rainfall relationship. This can be due to a set of differences in the two papers. Firstly, they use different rainfall data for their paper and control for contemporaneous, rather than lagged Monsoon rainfall. Secondly, specifications are estimated on a longer panel for which TRMM rainfall data is not available. This however, however, comes at the cost of loosing spatial variation in the rainfall measure as balancing the panel requires merging of districts to reflect district boundary changes. Last, but not least, the differences could also be due to the fact that they use different sets of fixed effects.

[43]Unfortunately, one can not "clean" the data by removing crime reported in cities falling into a district since the threshold city size was later changed to include only cities with at least 1 million inhabitants.

[44]There is some evidence that this is the case. An analysis for the 75 cities that report crime data for the period under study suggests an insignificant relationship between lagged Monsoon rainfall and violent crime, but a significant and positive relationship for property crimes. These results are available upon request from the author.

study which NREGA implementation phases seem to drive the estimated effect. The expectation is that this comes from districts that received earlier program support, as in general they are more vulnerable. I then highlight that the estimated effects are strongly driven by the districts that are considered by the Ministry of Home Affairs to be affected by left-wing extremism. Last, I study who is a target of violence and how this changes with the introduction of NREGA.

5.5 Heterogeneity

Effects by Implementation Phase NREGA was introduced in three distinct phases. This allows me to estimate the effects by NREGA implementation phase. This becomes insightful in determining which districts seem to be driving the overall observed effects. I present the results as before in that I study the relationship between a lagged monsoon rainfall and conflict over time. The results are presented in Figure 4.

The key observation is that the moderation of the monsoon rainfall-conflict relationship is driven by the districts that received NREGA in earlier phases. The estimated coefficient on monsoon rainfall is negative before NREGA was introduced and becomes generally insignificant for the period after NREGA was introduced. This pattern is visible for the districts that received NREGA support in phases 1 and 2, while for the districts in phase 3, there is no statistically discernible effect. The observed patterns are quite reasonable as the districts for phases 1 and 2 where poorer relative to the rest of India. The districts in earlier phases are the most vulnerable. For districts that received NREGA support in the third phase, the conflict-incidence regressions suggest a statistically insignificant relationship, while the conflict-intensity results suggest a mixed result. The monsoon rainfall and conflict coefficients are negative and significant at about the time of NREGA introduction date. But these coefficients become insignificant towards the end of the sample period. This highlights the fact that these poorer and more vulnerable districts drive this effect. The results are less conclusive for districts that are richer, on average.

As has been pointed out in the context discussion, NREGA was more likely to have been introduced early in the districts that were under the influence of Maoists. This makes it reasonable to study the distinct effect of NREGA in the specific districts that have been classified as being under the influence of left-wing extremism: districts that form the Red Corridor.

Effects by Conflict The rollout of NREGA suggests that districts that are under the influence of left-wing extremism were more likely to receive program support in the earlier phases. This makes it important to study the impact of the scheme on the relationship between the monsoon rainfall and conflict in districts that were under the influence of left-wing extremism, relative to the rest of India. I present the results by plotting out the estimated coefficients of monsoon rainfall over time. I constrain the analysis to the districts that were classified as being under the influence of left-wing extremism. I then do a separate analysis for the rest of India. The results are presented in Figure 5.

The top panel of Figure 5 shows the results for districts under left-wing extremist influence. These account for 130 of the 222 districts that have experienced variations in conflict over time. The coefficient pattern strongly follows the suggested pattern, indicating that monsoon rainfall was a predictor of conflict in these districts before NREGA was introduced there. This relationship has become a lot weaker since the introduction of NREGA. For non-left-wing-extremist affected districts that experienced conflict, the results are less clear (see bottom panel). The main variation, here, comes from Assam and Manipur, which account for 41% of all conflict events in the data. It is not clear whether one should separately study the conflict that arose from the Maoist insurgency in these states as, at least in Assam, there is a significant Maoist presence. To begin with, the left panel indicates that there is no relationship for conflict incidence: this is not surprising as there is hardly a district in these states where conflict is not experienced. For conflict intensity, the coefficients are negative at around the time of NREGA introduction, but it then becomes insignificant in more recent years.

An analysis of the agricultural output relationship for the Northeast is less straightforward as non-linearities in the monsoon rainfall-agricultural output relationship are much more pronounced there. A logarithmic transformation may not do justice to the agricultural-output relationship. Unfortunately, the production data for the Northeast is particularly thin, making it difficult to dig deeper into the underlying nature of the relationship. The results from this analysis suggest that the bulk of the NREGA effect comes from districts that are classified by the Ministry of Home Affairs as having been under the influence of left-wing extremism.

The next section provides evidence that less (targeted) violence against civilians drives moderation in NREGA support. This suggests that NREGA could help take civilians out of the line of fire.

NREGA and Targets of Violence The relationship between monsoon rainfall and conflict could be heterogeneous depending on who is the subject of violence. Vanden Eynde (2011) argues that civilians who face an income shock find themselves torn between becoming paid police informers or not. This option comes at a cost as insurgents react with more violence against these civilians. As NREGA primarily stabilizes rural incomes, there may be less targeted violence against civilians following adverse shocks. The conflict data allows for the subjects of violent activities to be a roughly classified into such groups as civilians, security forces, and terrorists. In Fetzer (2013), I highlight how this is done for most cases using trained machine learning algorithms, while relying on of humans, to classify ambiguous cases. I proceed as before, except that now I change the dependent variable to the number of conflict events in a year, where the subject of the event has been classified to be either a civilian, a member of the security forces or an insurgent. The results are presented in table 8.

Columns (1) to (3) of table 8 perform the analysis of the NREGA effect in regard to conflict incidence. Columns (4) to (6) study intensity. The coefficient pattern that emerges shows that all types of violence are responsive to lagged monsoon. However, in columns (1) and (4) the moderating effect of NREGA is most strongly seen for violence targeted against civilians. Violence against security forces is shown in columns (2) and (4), which also exhibit an NREGA effect. The sum of the two coefficients is actually positive but insignificant. The third column looks at incidences where a terrorist was the subject of the incidence of violence. Here, the moderating effect of NREGA appears to be weak. This evidence suggests that NREGA moderates the relationship between monsoon rainfall and violence, with the bulk of that effect coming from the reduced violence against civilians. This indicates that NREGA may help bring civilians out of the line of fire.

5.6 Other explanations

The key concern for identification is whether the timing of the NREGA treatment was correlated with other policies or events that could also have been correlated with monsoon rainfall and, through that, have an affect on the relationship between monsoon rainfall and conflict. In Appendix A.1, I rule out a whole range of alternative explanations for the observed moderation in the relationship between monsoon rainfall and conflict. The first set rules out other development schemes that could either directly or indirectly moderate the relationship between these two factors. This includes the Pradhan Mantri Gram Sadak Yojana Rural Road Construction Scheme (PMGSY), and the Integrated Action Plan, the latter which channels additional funds into districts affected by left-wing extremism. The second concern relates to a major military intervention'Operation Green Hunt". While officially not acknowledged this operation has been under way, the earliest, since late 2009. Due to a lack of data on troop deployment, I cannot directly address this issue but try to provide indirect evidence that it is not driving my results. Lastly, I rule out that NREGA does not capture the structurally different monsoon rainfall and conflict relationship that is due to the mining activity in a district. This could arise since the timing of the introduction of NREGA is correlated with a commodity price boom. Lastly, there could be an indirect way through which road construction under NREGA affects conflict. Fearon and Laitin (2003) highlight that guerrilla warfare thrives in places the state finds difficult to access. Insurgency movements may have an incentive to prevent public infrastructure development, in particular, roads in rural areas as this could lead to more government presence. That is to say, road construction could trigger more conflict. If road construction under NREGA was correlated with a lagged monsoon season rainfall, this could explain part of the observed effects. In Appendix A.3, I provide evidence that NREGA road construction may be correlated with conflict levels; I show that this effect is unlikely to be associated with monsoon rainfall.

Thus far, this paper has focused on how NREGA has led to an inward rotation of the monsoon rainfall-conflict relationship. The implicit argument is that NREGA provides insurance against adverse income shocks. If the relationship between local monsoon shocks and conflict is income, then insulating income from adverse monsoon shocks should break this relationship. The next section answers the question as to whether NREGA provides insurance. It also provides a rough quantification exercise for determining how much insurance is provided.

6 Does NREGA Provide Insurance?

The findings of this research show that the introduction of NREGA and the income supports it provides to vulnerable communities has changed the relationship between monsoon shocks and conflict. The implicit argument is that NREGA provides the kind of insurance program that mitigates the impact of adverse weather shocks on rural household incomes by providing a cushion and, thus, breaks the link between income and conflict and some forms of crime. This relationship needs to be empirically verified through a study of how NREGA participation and program expenditures respond to the adverse shocks posed by a deficient monsoon. The hypothesis is that NREGA take-up is responsive to adverse monsoon shocks: the slope that links monsoon rainfall and NREGA participation is downward sloping. This indicates that positive rainfalls translate into lower NREGA participation as, in these situations, the workers' outside options are better, which makes minimum-wage employment under NREGA less appealing.

I also study how adverse monsoon shocks in the preceding growing season translate into increased NREGA participation. Since the harvest season is towards the end of the year (November/December), we should see a pronounced NREGA employment response in the following spring, before the planting season. The results from the baseline analysis are presented in Table 9. The first column measures the log of the total expenditures on ongoing projects in a fiscal year. The elasticity is negative, which indicates that a good monsoon rainfall in the preceding growing season translates into low expenditures. A 1% decrease in a monsoon rainfall increases expenditures by 0.257%.

Columns (2) to (4) study the margins of participation. The overall take-up is measured as the total number of days worked, under the scheme, in a district and year. Again, the coefficient is negative: With a good monsoon, there is less need for NREGA employment. The coefficient indicates that a 1% drop in the monsoon rainfall in the preceding growing season increases NREGA participation by 0.21%. This is high, but not unreasonable, given the large share of self-employed farmers in rural India. The overall take-up effect is decomposed into extensive- and intensive-margin participation in columns (3) and (4). The extensive margin measures the share of household participation; the intensive margin is the log of the number of days worked per household. Since the program provides assistance on a per-household basis, this is the correct way to measure extensive-margin participation. This measure also indicates a negative relationship. The intensive-margin coefficient suggests that a 1% decrease in the monsoon rainfall, increases the number of days worked under NREGA by 0.12%. This suggests that a significant share of the overall observed participation response in column (2) is driven by extensive-margin participation.

Columns (5) and (6) consider the heterogeneity in extensive-margin participation: the relationship between the lagged monsoon rainfall and NREGA participation by implementation phase (column (5) and by whether a district is classified as being under left-wing extremist influence, according to the Ministry of Home Affairs. The pattern suggests that the effect of a monsoon rainfall on extensive-margin participation is strongest for the districts in phase 1. Furthermore, in districts where left-wing extremism pervades, the relationship between NREGA take-up and the monsoon rainfall is also a lot stronger. This corresponds well with the findings in the previous section. The moderation noted in the findings was mostly driven by the districts in the first phase and the districts that were vulnerable to left-wing extremist activity.[45]

Anecdotal accounts suggest that Naxalites encourage households to demand NREGA employment. In some areas of Jharkhand, local Naxal units put up posters urging villagers to claim their right to employment. "The Naxalites do not oppose NREGA. They ask us to work and demand employment"[46] The Council for Social Development collected evidence suggesting further that the Naxals were not blocking any activities that were being implemented under the NREG Scheme. The report states that as a combined result of the NREGA and the Naxalites' pressure, contractors are paying higher wages to manual workers in the areas hit by left-wing extremism. There is however, some evidence that Maoists are blocking road and bridge construction under NREGA.

These findings suggests that NREGA does function as an insurance program: NREGA employment and program participation is higher following a local monsoon shock. This effect is driven by the districts that are most vulnerable to conflict: the districts that received NREGA support in the earlier phases of the program and the districts that were classified as being under the influence of left-wing extremism. With the above in mind, the open question becomes: *How much insurance is provided?* To answer this, it is necessary to determine to what extent NREGA expenditures offset the income losses incurred as a result of adverse monsoon shocks. Ideally this question would be addressed by using a household panel dataset. To date, such a dataset that covers many parts of India does not exist.[47] In this paper, the best I can do is exploit the district-level variation in agricultural output value per capita that are due to monsoon variations. This will enable me to arrive at a rough quantification of how much insurance is provided. I relate this variation with NREGA expenditures per district. This allows for a crude quantification exercise that measures the extent to which district output losses are compensated

[45] Similar to the previous analysis, I also present results from a non-parametric approach to highlight potential non-linearities in the response of NREGA participation with regards to Monsoon rainfall. These are presented in Figure 8.

[46] See Hindustan Times, 22.03.2008.

[47] The India Human Development Survey is a candidate dataset; the first round of interviews was completed in 2005 and the second round of data was collected in 2011-2012. Unfortunately, the data are not released until early 2015.

by increased NREGA expenditures. The question is by how much an INR 100 loss in agricultural output per capita is offset by an increased NREGA per capita expenditure. A study of the relationships in the levels of expenditure and agricultural output per capita allows a direct comparison of these two factors. As there are significant outliers in both sets of data, I trim the bottom and top 1% of observations from both variables.[48]

The results are presented in Table 10. Column (1) is the agricultural production function that links monsoon rainfall with the output value in the levels of expenditure and agricultural output per capita. The coefficient on monsoon rainfall shows a semi-elasticity, indicating that a 10% increase in the monsoon rainfall increases the nominal agricultural-output value by INR 54.1. This can be interpreted as the first stage of the analysis. Column (2) estimates how a lagged monsoon rainfall translates into an increased NREGA per-district expenditure. The coefficient indicates that a 10% reduction in a monsoon rainfall in the preceding growing season increases the per capita NREGA expenditure by INR 10.1. This suggests that there is partial insurance coverage: every income loss of INR 10 is accommodated by an increase in NREGA expenditure by INR 1.86. This can be further refined. In column (3) I present the results from an instrumental variables exercise. Here, lagged agricultural output is instrumented with lagged monsoon rainfall. The coefficient suggests that an INR 100 loss in agricultural output due to monsoon variation translates into an increased NREGA expenditure of INR 30.1. If we study labor expenses only, the coverage is INR 21.6. The combined results suggest that NREGA expenditures, in a district, offset a significant share of the risk crop cultivation faces as a result of a local weather variation.

Nevertheless, these estimates should be taken with a grain of salt. First, the number is not adjusted for any marginal leakage in the program. Niehaus and Sukhtankar (2013a,b) found that leakage rates from the NREGA program were significant, particularly during the years it was first introduced. This would suggest that the NREGA expenditure measure is likely to be an upper bound. On the other hand, the results in the previous section suggest that there is significant indirect insurance due to the stabilization of agricultural wages. This has an indirect insurance effect that is not captured by a mere accounting of NREGA expenditures that flow into a district.

7 Conclusion

This paper has studied the impact of a social insurance program on conflict in India. The existing conflict literature has devised various identification strategies to arguably exploit exogenous variations in incomes to study the relationship between income and conflict. The findings of this literature have a direct policy implication: any measure that helps insulate household incomes from adverse shocks should moderate the relationship between these exogenous productivity shocks and conflict. This paper has

[48]Please consult appendix B.5 for a discussion of how agricultural output value is scaled to ensure that it comes closer to the true agricultural output using district domestic product for 2000.

taken up this question and evaluates the impact of the introduction of a public employment program established through the National Rural Employment Guarantee Act in India on the relationship between local monsoon shocks and conflict and crime.

The key findings suggest that the introduction of this public employment program has eliminated the link between monsoon shocks and conflict and some forms of crime. But productivity shocks continue to affect rural areas, even after the introduction of the public employment program. Nevertheless, these shocks have ceased to translate into conflict as a result of having the public employment program in place. As conflict is a phenomena that has a lot of persistence, removing the link between productivity shocks and conflict can lead to persistently lower levels of conflict. This is the case in India. The insurance value delivered by the public employment scheme is significant. A simple quantification exercise suggests that roughly one third of district-level income losses due to adverse monsoon conditions are directly offset by the increased expenditures the program provides. This only captures the direct transfers. An added indirect insurance benefit comes in the reduction of productivity shocks on wages because of the sense of economic security the program creates.

The paper has important implications for policy makers. Incomes in developing countries are much more volatile. This leaves households exposed to a lot more risk in comparison to developed countries. Climate models suggest that erratic weather events could become even more pronounced. This has led to concerns about future increasing conflict (see Hsiang et al., 2013). Yet, many developing countries have not been able to devise policies to provide adequate protection. This paper highlights that social insurance that takes the form of a public employment program may be a policy that can also be adopted in other developing countries. The appealing features of a public employment program are that it induces the kind of self-targeting that can help overcome the general adverse selection issues that are common to transfer schemes. The indirect benefits derived from the lower likelihood of conflict and the reduced intensity of conflict are important margins through which social insurance can provide social benefits.

References

Abadie, A. (2005). Semiparametric Difference-in-difference Estimators. *The Review of Economic Studies 72*(1), 1–19.

Abadie, A. and J. Gardeazabal (2008). Terrorism and the world economy. *European Economic Review 52*(1), 1–27.

Aggarwal, S. (2014). Do Rural Roads Create Pathways Out of Poverty? Evidence from India. *mimeo*.

Akresh, R. and D. D. Walque (2008). Armed conflict and schooling: Evidence from the 1994 Rwandan genocide. *mimeo* (April).

Annan, J. and C. Blattman (2014). Can employment reduce lawlessness and rebellion? Experimental evidence from an agricultural Intervention in a fragile state. *mimeo*.

Azam, M. (2011). The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment. *IZA Discussion Paper 6548*.

Banerjee, K. and P. Saha (2010). The NREGA, the Maoists and the developmental woes of the Indian state. *Economic and Political Weekly xlv*(28), 42–48.

Bazzi, S. and C. Blattman (2014). Economic Shocks and Conflict: Evidence from Commodity Prices. *American Economic Journal: Macroeconomics 6*(4), 1–38.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra (2012). Can Rural Public Works Affect Agricultural Wages? Evidence from India. *CSAE Working Paper*.

Besley, T. and S. Coate (1992). Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review 82*(1), 249–261.

Besley, T., T. Fetzer, and H. Mueller (2014). The Welfare Cost of Lawlessness: Evidence from Somali Piracy. *forthcoming, Journal of the European Economics Association*.

Besley, T. and T. Persson (2008). The Incidence of Civil War: Theory and Evidence. *NBER Working Paper*.

Blakeslee, D. and R. Fishman (2014). Weather Shocks, Crime, and Agriculture: Evidence from India. *mimeo*.

Blattman, C. and J. Annan (2010). The Consequences of Child Soldiering. *Review of Economics and Statistics 92*(4), 882–898.

Blattman, C., N. Fiala, and S. Martinez (2014). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *The Quarterly Journal of Economics 129*(2), 697–752.

Blattman, C. and E. Miguel (2010). Civil War. *Journal of Economic Literature 48*(1), 3–57.

Buhaug, H. and G. Wischnath (2014). Rice or riots: On food production and conflict severity across India. *in Press, Political Geography*.

Burgess, R., O. Deschenes, D. Donaldson, and M. Greenstone (2011). Weather and Death in India. Technical report, MIT.

Burgess, R. and D. Donaldson (2010). Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era. *American Economic Review: Papers & Proceedings 100*(2), 449–453.

Burnicki, A. C., D. G. Brown, and P. Goovaerts (2007). Simulating error propagation in land-cover change analysis: The implications of temporal dependence. *Computers, Environment and Urban Systems 31*(3), 282–302.

Chassang, S. and G. P. i. Miquel (2009). Economic Shocks and Civil War. *Quarterly Journal of Political Science 4*(3), 211–228.

Ciccone, A. (2011). Economic Shocks and Civil Conflict: A Comment. *American Economic Journal: Applied Economics 3*(4), 215–227.

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery (2010). Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics 5*(1), 104–135.

Cole, S., J. Tobacman, and P. Topalova (2008). Weather insurance: Managing risk through an innovative retail derivative. *mimeo*.

Collier, P. and A. Hoeffler (1998). On economic causes of civil war. *Oxford Economic Papers 50*(4), 563–573.

Collier, P. and A. Hoeffler (2004). Greed and grievance in civil war. *Oxford Economic Papers 56*(4), 563–595.

Collobert, R., J. Weston, and L. Bottou (2011). Natural language processing (almost) from scratch. *The Journal of Machine . . . 12*, 2461–2505.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

Dasgupta, A., K. Gawande, and D. Kapur (2014). Anti-Poverty Programs Can Reduce Violence: India's Rural Employment Guarantee and Maoist Conflict. *mimeo*.

Dash, S. P. (2006). *Naxal Movement and State Power: With Special Reference of Orissa* (1st editio ed.). New Delhi: Sarup & Sons.

Dee D. P., Uppala S. M., Simmons A. J., Berrisford P., Poli P., Kobayashi S., Andrae U., Balmaseda M. A., Balsamo G., Bauer P., Bechtold P., Beljaars A. C., L. M. van de Berg, Bidlot J., Bormann N., Delsol C., Dragani R., Fuentes M., Geer A. J., Haimberger L., Healy S. B., Hersbach H., Hólm E. V., Isaksen L., Kållberg P., Köhler M., Matricardi M., McNally A. P., Monge-Sanz B.M., Morcrette J.J., Park B.K., Peubey C., de Rosnay P., Tavolato C., J. N. Thépaut, and F. Vitart (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society 137*, 553–597.

Deshpande, R. S. and K. Shah (2010). *Agrarian Crisis and Farmer Suicides* (12 ed.). New Delhi: Sage Publications India Ltd.

Donaldson, D. (2013). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *forthcoming, American Economic Review*.

Dube, O. and J. F. Vargas (2013). Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies 80*(4), 1384–1421.

Duflo, E. and R. Pande (2007). Dams. *The Quarterly Journal of Economics 122*(2), 601–646.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review 97*(01), 75.

Ferrara, E. L. and M. Harari (2013). Conflict, Climate and Cells: A Disaggregated Analysis. *CEPR Discussion Paper 9277*.

Fetzer, T. R. (2013). Developing a Micro-Level Conflict Data Set for South Asia. *mimeo*.

Gadenne, L. (2014). Optimal non-linear commodity taxation in developing countries: Theory and an application to India. *mimeo*.

Gawande, K., D. Kapur, and S. Satyanath (2012). Renewable Resource Shocks and Conflict in India's Maoist Belt. *mimeo*.

Gentzkow, M. and J. M. Shapiro (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica 78*(1), 35–71.

Gomes, J. (2012). The Political Economy of the Maoist Conflict in India: An Empirical Analysis. *mimeo*.

Hegre, H. and N. Sambanis (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution 50*(4), 508–535.

Hoelscher, K., J. Miklian, and K. C. Vadlamannati (2012). Hearts and mines: A district-level analysis of the Maoist conflict in India. *International Area Studies Review 15*(2), 141–160.

Hornbeck, R. and P. Keskin (2014). The Historically Evolving Impact of the Ogallala Aquifer: Agricultural Adaptation to Groundwater and Drought †. *American Economic Journal: Applied Economics 6*(1), 190–219.

Hsiang, S. M., M. Burke, and E. Miguel (2013). Quantifying the influence of climate on human conflict. *Science 341*(6151), 1–7.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology 8*(1), 38–55.

Imbert, C. and J. Papp (2014). Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. *forthcoming American Economic Journal: Applied Economics*.

Iyengar, R., J. Monten, and M. Hanson (2011). Building Peace: The Impact of Aid on the Labor Market for Insurgents. *NBER Working Paper 17297*.

Iyer, L., A. Mani, P. Mishra, and P. Topalova (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied Economics 4*(4), 165–193.

Iyer, L. and P. Topalova (2014). Poverty and Crime : Evidence from Rainfall and Trade Shocks in India. *Harvard Business School Working Paper 14067*.

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

Jia, R. (2014). Weather Shocks, Sweet Potatoes and Peasant Revolts in Historical China. *The Economic Journal 124*(575), 92–118.

Katyal, R., M. Sardana, and D. Satyanarayana (2001). *Estimates of District Domestic Product*. New Delhi: Socio-Economic Research Centre.

Khanna, G. and L. Zimmermann (2013). Guns and Butter ? Fighting Violence with the Promise of Development. *mimeo*.

Kudamatsu, M., T. Persson, and D. Strömberg (2014). Weather and Infant Mortality in Africa. *mimeo*.

Kung, J. K.-s. and C. Ma (2014). Can cultural norms reduce conflicts? Confucianism and peasant rebellions in Qing China. *Journal of Development Economics 111*, 132–149.

Leon, G. (2009). Civil conflict and human capital accumulation: The long term effects of political violence in Peru. *mimeo*.

Leung, Y., Y. Ge, J. Ma, and J. Wang (2005). Measurement Errors and their Propagation in the Registration of Remote Sensing Images. In *Developments in Spatial Data Handling*, pp. 285–297. Springer Berlin Heidelberg.

Li, X.-H., Q. Zhang, and C.-Y. Xu (2012). Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *Journal of Hydrology 426-427*, 28–38.

Lilleor, H. and X. Giné (2005). Weather insurance in semi-arid India. *mimeo*.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy 112*(4), 725–753.

Ministry of Rural Development (2008). The National Rural Employment Guarantee Act 2005 (NREGA)-Operational Guidelines. Technical report, Ministry of Rural Development, Government of India, New Delhi.

Ministry of Rural Development (2009). Joint Convergence Guideline. Technical report, Ministry of Rural Development.

Ministry of Rural Development (2012). *MGNREGA Sameeksha, An Anthology of Research Studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005, 2006–2012*. New Delhi: Orient BlackSwan.

Morgan, R. and D. Reiter (2013). How War Makes the State: Insurgency, External Threat, and Road Construction in India. *mimeo*.

Morrison, B. C. (2012). Grievance, Mobilisation and State Response: An examination of the Naxalite Insurgency in India. *Journal of Conflict Transformation & Security 2*(1).

Nichols, A. and R. Zeckhauser (1982). Targeting Transfers through Restrictions on Recipients. *American Economic Review: Papers & Proceedings 72*(2), 372–377.

Niehaus, P. and S. Sukhtankar (2013a). Corruption Dynamics: The Golden Goose Effect. *American Economic Journal: Economic Policy 5*(4), 230–269.

Niehaus, P. and S. Sukhtankar (2013b). The marginal rate of corruption in public programs: Evidence from India. *Journal of Public Economics 104*, 52–64.

Pandita, R. (2011). *Hello Bastar: The Untold Story of India's Maoist Movement*. New Delhi: Tranquebar Press.

Planning Commission (2005). Report of the Inter Ministry Task Group on: Investment Credit and Technical Support to Promote Self-employment in Agriculture, Horticulture, Afforestation, Dairying and Agro-processing. Technical Report January, Planning Commission, New Delhi.

Planning Commission (2008). Development Challenges in Extremist Affected Areas. Technical report, Planning Commission, New Delhi.

Prasad, P. (1987). Agrarian Violence in Bihar. *Economic and Political Weekly 22*(22), 847–852.

Radhakrishna, R. and S. Ray (Eds.) (2006). *Handbook of Poverty in India: Perspectives, Policies, and Programmes*. Oxford: Oxford University Press.

Rahman, H. and D. Sengupta (2007). Preliminary Comparison of Daily Rainfall from Satellites and Indian Gauge Data. *mimeo*.

Ramana, P. V. (2007). *The Naxal Challenge: Causes, Linkages, and Policy Options* (1st editio ed.). Chennai: Pearson.

Rana, S. A. (2013). Terror financing or labour supply: the dominant channel of terrorist attacks in pakistan. *mimeo*, 1–51.

Ranjan, R. and V. Prasad (2012). Exploring Social Resilience in State Fragility: A Climate Change Perspective. In M. Hamza and C. Corendea (Eds.), *Climate Change and Fragile States: Rethinking Adaptation*, Number 16, Chapter 2. Bonn: United Nations University.

Rosenzweig, M. and H. Binswanger (1993). Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments. *Economic Journal 103*(416), 56–78.

Sarsons, H. (2011). Rainfall and Conflict. *mimeo*.

Schneider, U., A. Becker, P. Finger, B. Meyer-Christoffer, Anja; Rudolf, and M. Ziese (2011). GPCC Full Data Reanalysis Version 6.0 at 0.5 Degree: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data.

Sekhri, S. and A. Storeygard (2013). Dowry Deaths: Response to Weather Variability in India. *forthcoming, Journal of Development Economics*.

Shapiro, J. N., M. Callen, E. Berman, and J. Felter (2011). Do Working Men Rebel? Insurgency and Unemployment in Afghanistan, Iraq, and the Philippines. *Journal of Conflict Resolution 55*(4), 496–528.

Shrivastava, A. (2014). Civil conflict with rising wages and increasing state capacity: The Maoist insurgency in India. *mimeo*.

Silva, J. M. C. S. (2011). poisson : Some convergence issues. *mimeo*, 1–8.

Silva, J. M. C. S. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(November), 641–658.

Singh, P. (2013). Impact of Terrorism on Investment Decisions of Farmers: Evidence from the Punjab Insurgency. *Journal of Conflict Resolution*.

Smith, S. R., M. A. Bourassa, and M. Long (2011). Pirate attacks affect Indian Ocean climate research. *Eos, Transactions American Geophysical Union 92*(27), 225.

Vanden Eynde, O. (2011). Targets of Violence: Evidence from India's Naxalite Conflict. *mimeo*.

Verma, S. (2011). Far Reaching Consequences of the Naxalite Problem in India. Technical Report July, Rakshak Foundation.

Voors, M. J., E. E. M. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. V. Soest (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review 102*(2), 941–964.

Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: evidence from the Indian Employment Guarantee Scheme. *mimeo*.

Figures for Main Text

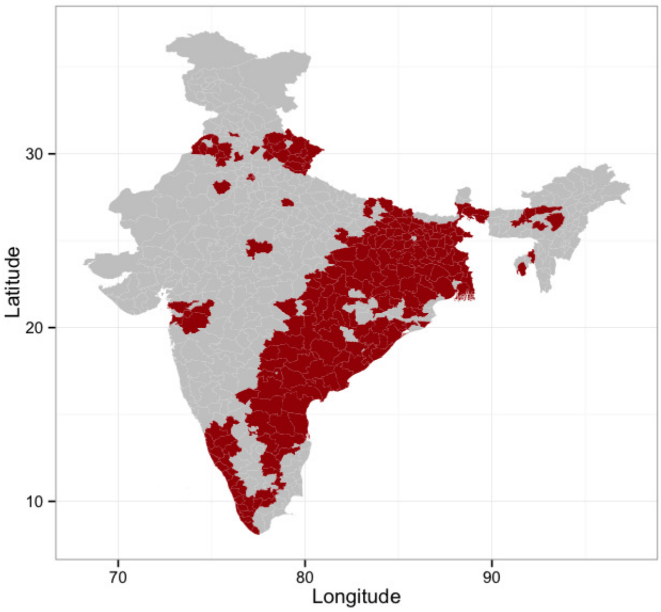

Figure 1: Districts Affected by Left-Wing Extremism According to Government of India.

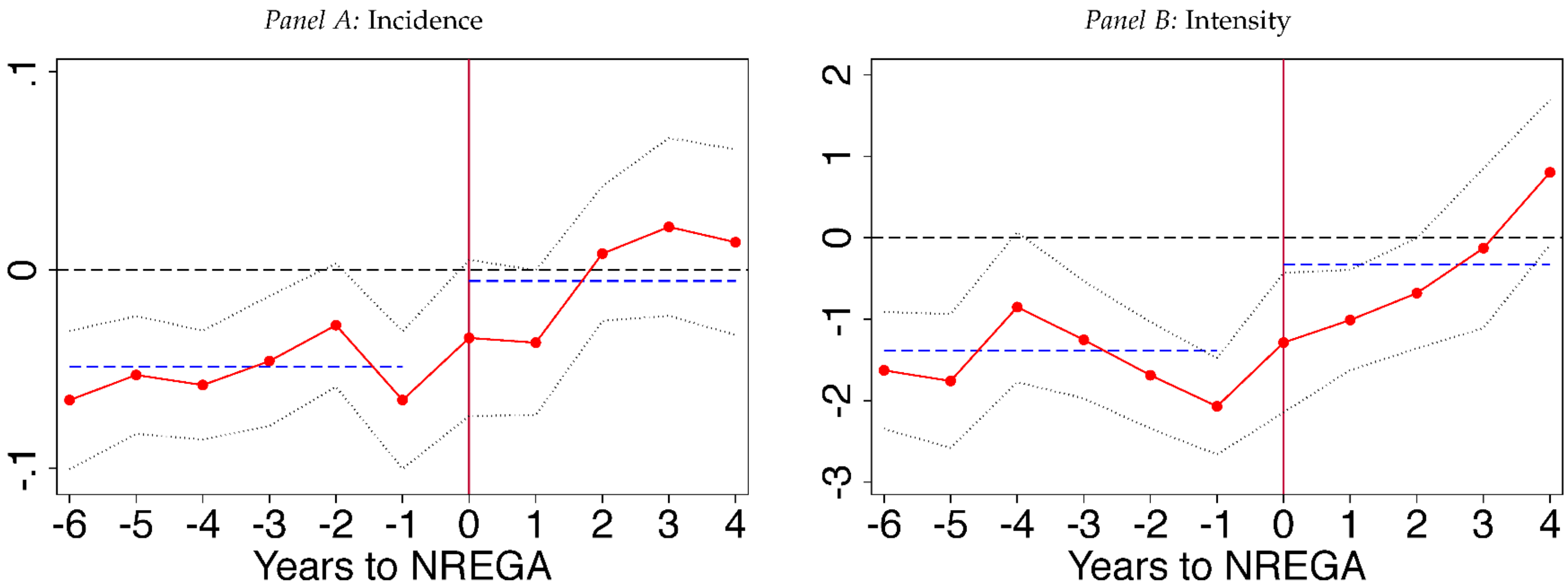

Figure 2: Effect of Monsoon Rain on Conflict Over Time. The vertical line indicates the NREGA introduction date. The blue dashed lines indicate the coefficients obtained from a simple regression interaction lagged Monsoon rainfall with the NREGA treatment indicator. The red line are each point estimates of the relationship between lagged Monsoon rainfall and conflict. 95% confidence bands are indicated as dotted black lines.

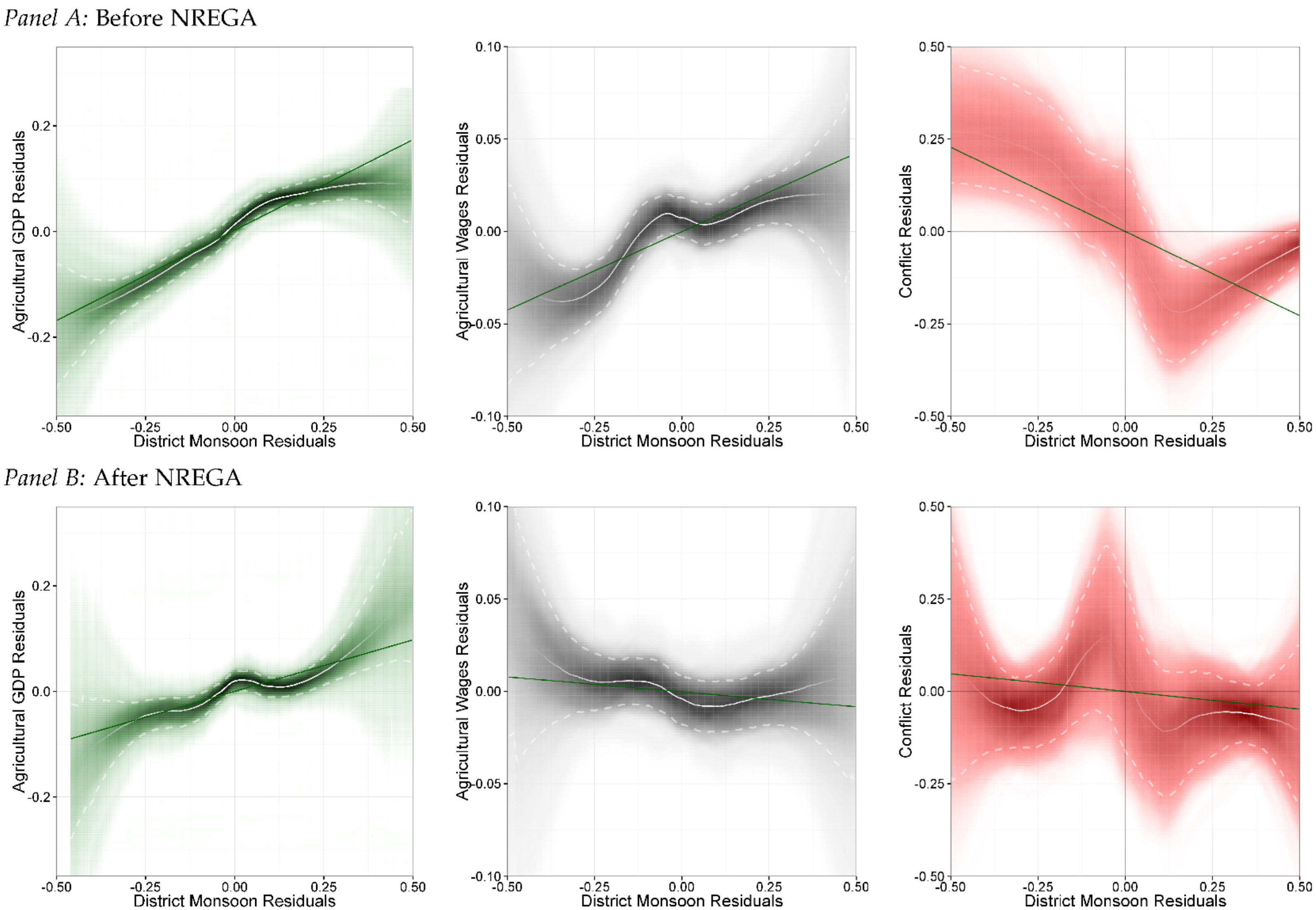

Figure 3: Non-Parametric Watercolor Regressions as in Hsiang et al. (2013): Effect of Monsoon Rain on agricultural output per Capita, Wages and Conflict Before and After Introduction of NREGA. 95% confidence bands are indicated as dashed lines. The color shading is related to the overall density of Monsoon rainfall realizations along the horizontal axis and to the density of fitted values from loess regressions along the vertical axis.

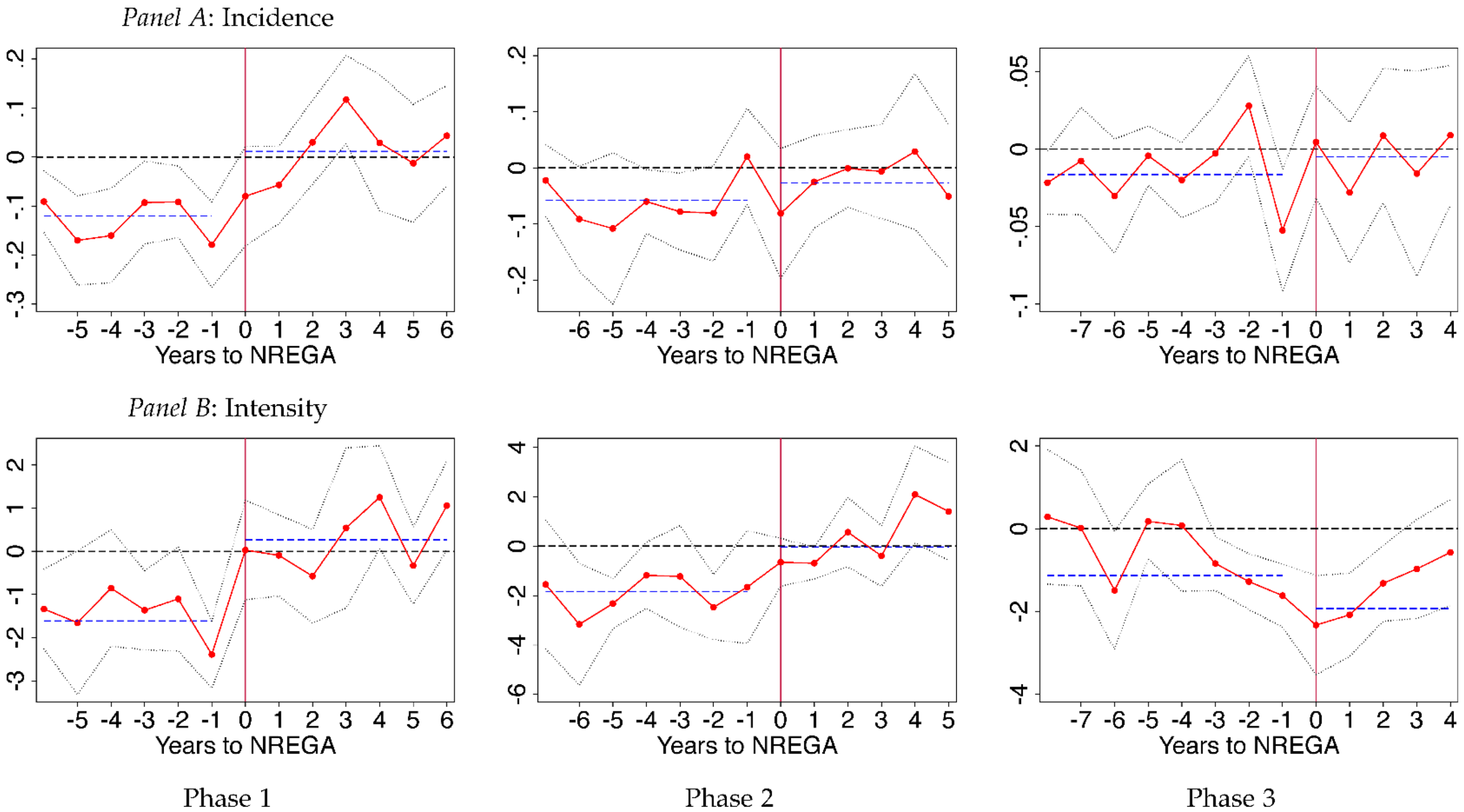

Figure 4: Effect of Monsoon Rain on Conflict Over Time By NREGA Implementation Phase. The vertical line indicates the NREGA introduction date. The blue dashed lines indicate the coefficients obtained from a regression interaction lagged Monsoon rainfall with the NREGA treatment indicator. The red line are each point estimates of the relationship between lagged Monsoon rainfall and conflict. 95% confidence bands are indicated as dotted black lines.

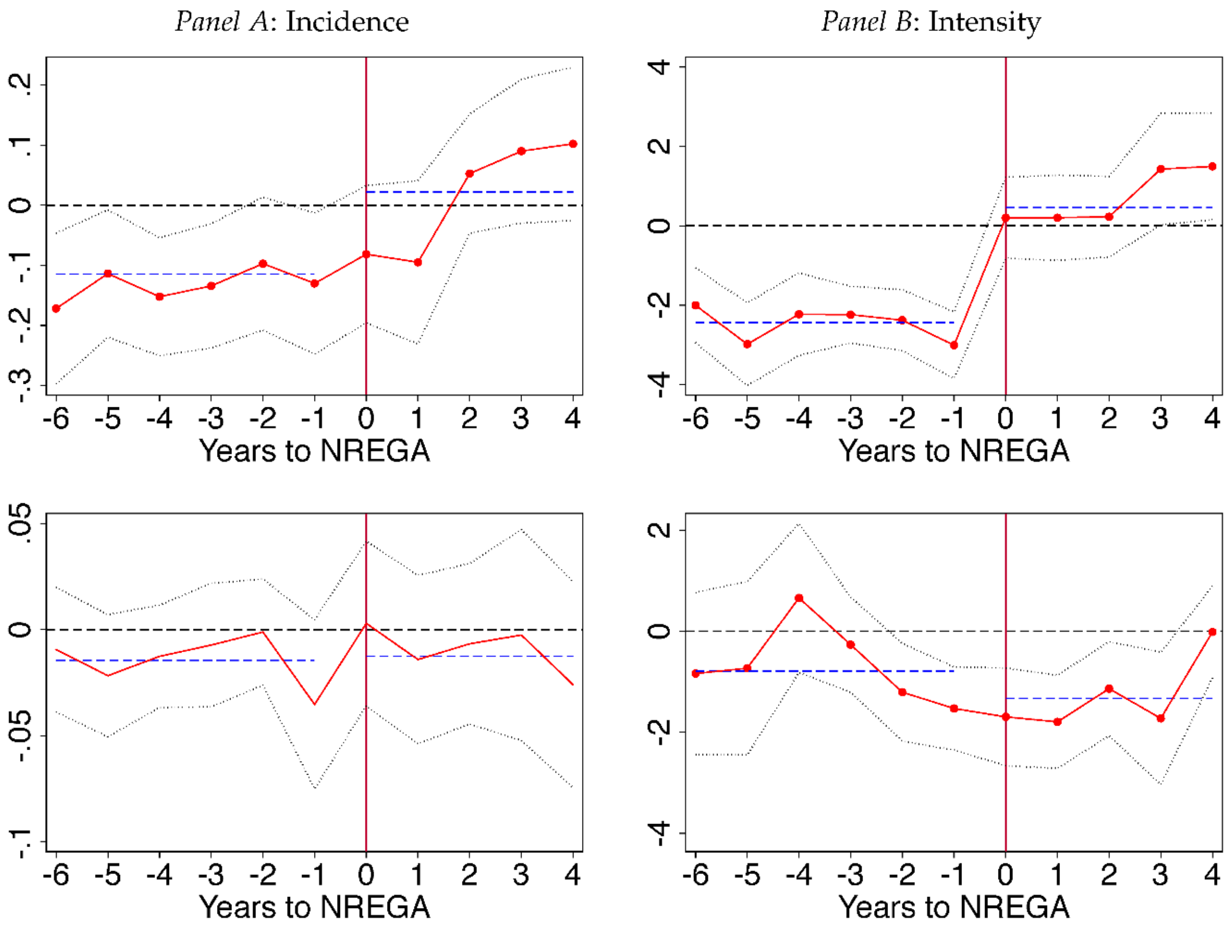

Figure 5: Effect of Monsoon Rain on Conflict Over Time in Red Corridor (top) and the Rest of India (bottom). The vertical line indicates the NREGA introduction date. The blue dashed lines indicate the coefficients obtained from a regression interaction lagged Monsoon rainfall with the NREGA treatment indicator. The red line are each point estimates of the relationship between lagged Monsoon rainfall and conflict. 95% confidence bands are indicated as dotted black lines.

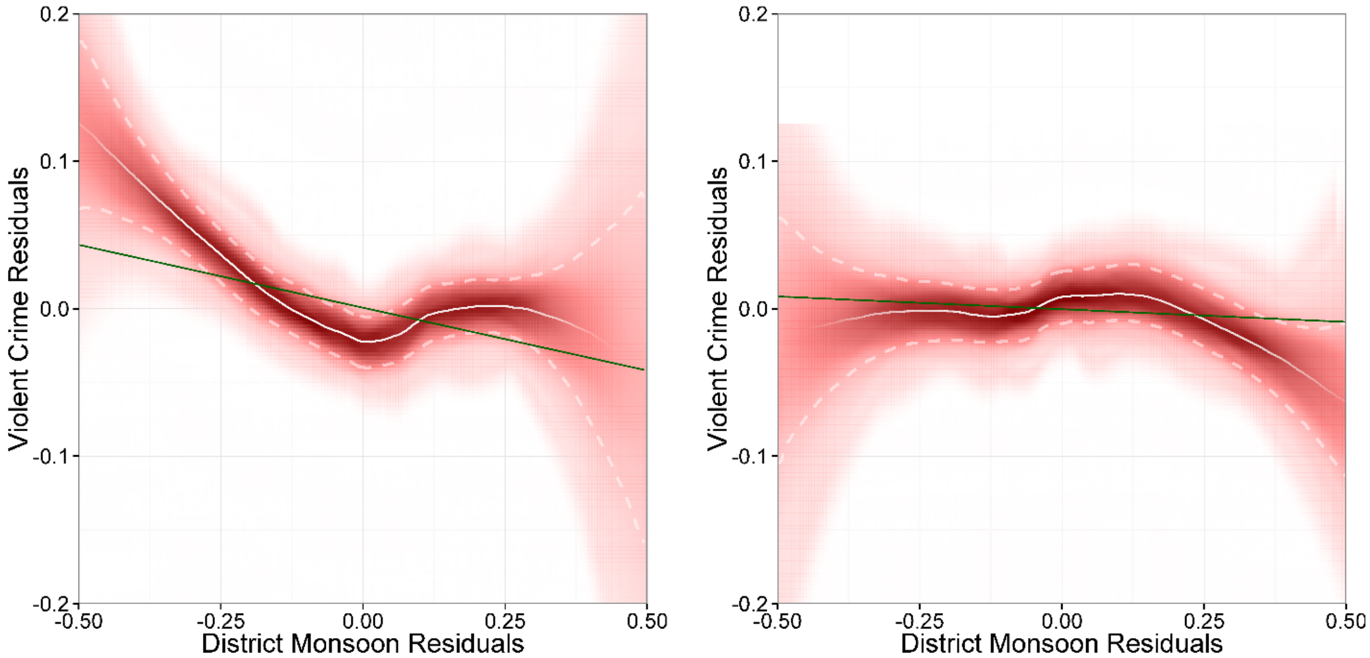

Figure 6: Non-Parametric Watercolor Regressions as in Hsiang et al. (2013): Relationship between lagged Monsoon rainfall and violent crime before (left) and after (right) the introduction of NREGA. 95% confidence bands are indicated as dashed lines. The color shading is related to the overall density of Monsoon rainfall realizations along the horizontal axis and to the density of fitted values from loess regressions along the vertical axis.

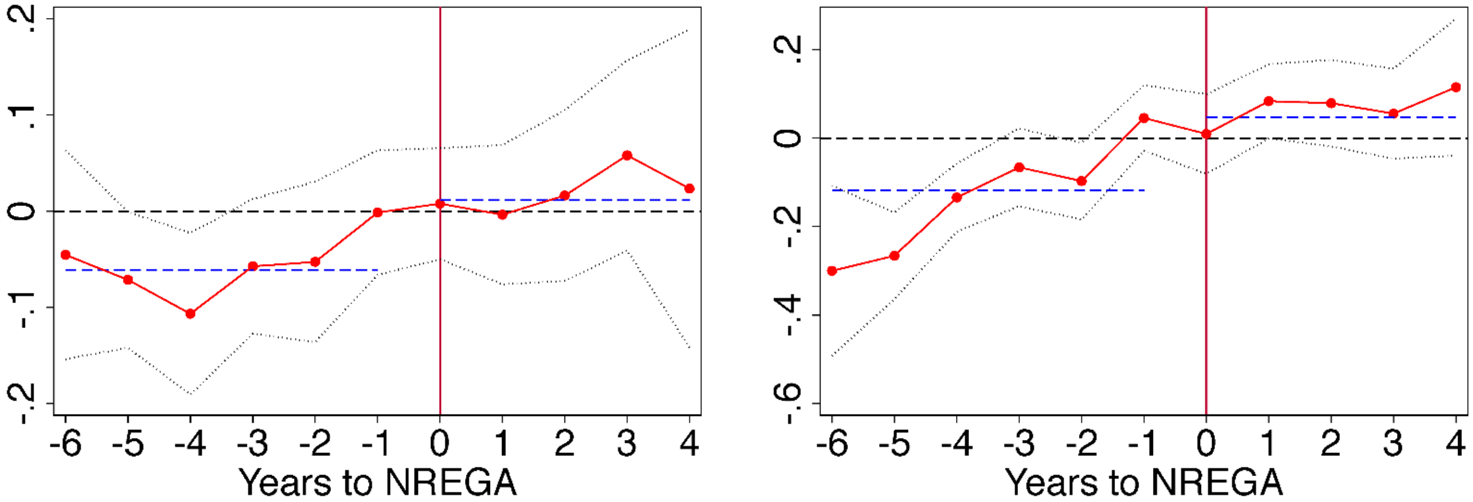

Figure 7: Effect of Monsoon Rain on Violent Crime (left) and on Crimes Against Public Order (right) over Time. The vertical line indicates the NREGA introduction date. The blue dashed lines indicate the coefficients obtained from a regression interaction lagged Monsoon rainfall with the NREGA treatment indicator. The red line are each point estimates of the relationship between lagged Monsoon rainfall and conflict. 95% confidence bands are indicated as dotted black lines.

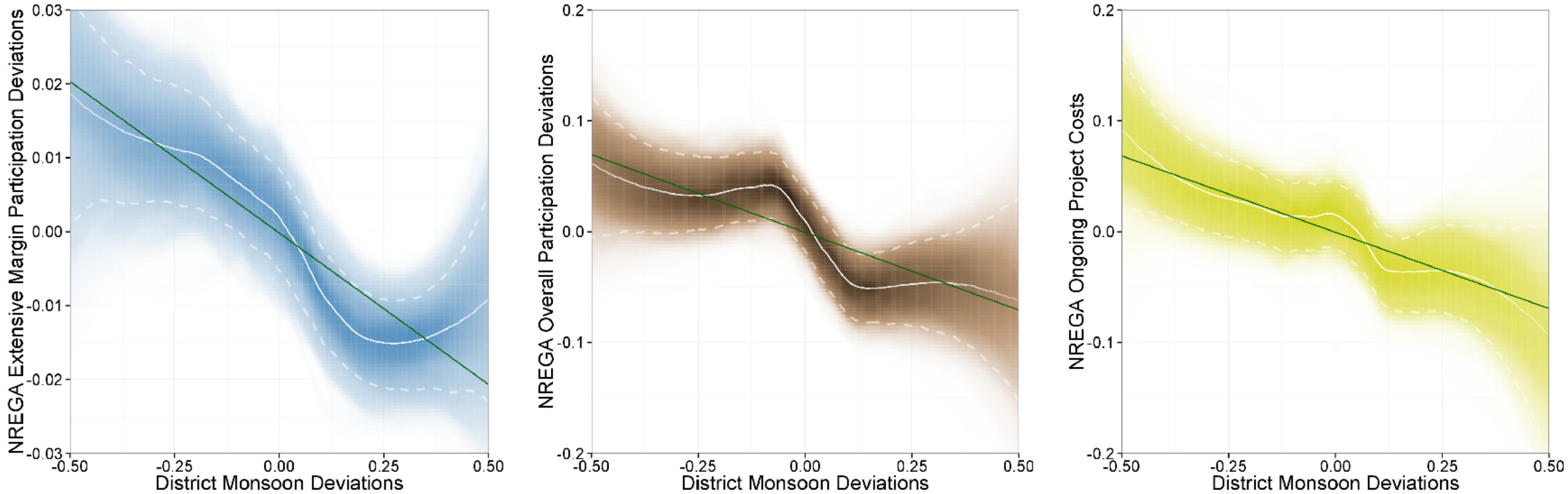

Figure 8: Non-Parametric Watercolor Regressions as in Hsiang et al. (2013): NREGA Take-Up and Lagged Monsoon Rainfall. 95% confidence bands are indicated as dashed lines. The color shading is related to the overall density of Monsoon rainfall realizations along the horizontal axis and to the density of fitted values from loess regressions along the vertical axis.

Tables for the Main Text

Table 1: Summary Statistics and Socio-Economic Characteristics of Districts Before NREGA by NREGA Implementation Phase

		Phase 1	Phase 2	Phase 3
Conflict	Left-Wing Affected	0.563	0.372	0.233
	Any Violence	0.270	0.166	0.105
	Conflict Events	1.643	1.001	0.634
Income	Agricultural Output Value per Capita [INR]	2746.668	2962.016	4792.635
	Agricultural Wages [INR]	53.628	62.165	77.525
	Share of District Night Lights	0.383	0.463	0.666
Weather	Average Monsoon Season Temp [Degrees]	23.753	24.401	23.610
	Annual Rainfall [mm]	1333.619	1446.971	1258.493
	Monsoon Season Rainfall [mm]	1028.777	1052.379	878.994
	NDVI Index	0.483	0.512	0.491
Terrain	Elevation	476.021	415.810	418.082
	Ruggedness	47.760	54.268	67.933
Socio-Economic	Rural Population [share]	0.853	0.808	0.715
	Tribal Population [share]	0.226	0.163	0.112
	Scheduled Caste [share]	0.154	0.151	0.149
	Illiterate Population [share]	0.525	0.472	0.414
	Household Size [persons]	5.400	5.515	5.414
	Population younger than 6 [share]	0.262	0.253	0.239
	Population Growth 1991-2001	21.652	24.445	21.213
	Gender Gap [per 1000 inhabitants]	25.114	21.568	20.414
Infrastructure	Primary School [share]	0.810	0.820	0.857
	Mudroad [share]	0.679	0.657	0.575
	Permanent Housing [share]	0.356	0.434	0.566
	Primary Health Care [share]	0.322	0.374	0.457
	Electricity [share]	0.678	0.784	0.909
	Bus Stop [share]	0.329	0.401	0.561
	Post Office [share]	0.368	0.467	0.601
NREGA	Expenditure per Capita [INR]	436.936	523.060	247.880
	Labor Expenditure per Capita [INR]	302.197	358.869	183.402
	Days per Household	47.974	42.496	40.486
	Share of Households Participating	0.442	0.382	0.204

Notes: Socio-economic and district Infrastructure statistics based on the 2001 Census for India. Infrastructure statistics is the share of villages in a district with access to a particular type of infrastructure.

Table 2: Before the Introduction of NREGA: Reduced Form Relationship between Rainfall, Agricultural Production, Wages and Violence

	Agricultural Income		Conflict	
	(1)	(2)	(3)	(4)
	ln(Output/Capita)	ln(Wage)	Incidence	Intensity
log(Monsoon)	0.362***	0.058***	-0.030**	-0.897***
	(0.086)	(0.018)	(0.015)	(0.309)
Observations	3239	1419	3843	932
Number of Districts	471	314	543	144
Estimation	OLS	OLS	OLS	Poisson

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects and constrain the analysis to the period before NREGA was introduced. Columns (1) and (2) study agricultural production and wages on an unbalanced annual district level panel, using contemporaneous Monsoon rainfall as independent variable. Column (2) also controls for state- by NREGA implementation phase linear time trends. Columns (3) and (4) are estimated on a balanced district level annual panel. Column (3) is a linear probability model using a dummy variable as dependent variable indicating whether a district experienced any conflict events in a given year. Column (4) estimates a Poisson regression with the dependent variable being the number of conflict events per district and year. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. For columns (1)-(3) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: After the Introduction of NREGA: Reduced Form Relationship between Rainfall, Agricultural Production, Wages and Violence

	Agricultural Income		Conflict	
	(1)	(2)	(3)	(4)
	ln(Output/Capita)	ln(Wage)	Incidence	Intensity
log(Monsoon)	0.374***	0.062***	-0.049***	-1.386***
	(0.078)	(0.019)	(0.018)	(0.291)
NREGA x log(Monsoon)	-0.132	-0.086***	0.043***	1.058***
	(0.083)	(0.016)	(0.014)	(0.392)
F-Test	6.36	1.68	.06	2.02
p-value	.01	.19	.8	.16
Observations	4480	2455	7059	2760
Number of Districts	471	336	543	222
Estimation	OLS	OLS	OLS	Poisson

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects. Columns (1) and (2) study agricultural production and wages on an unbalanced annual district level panel from 2000-2009 and 2001-2010 respectively, using contemporaneous Monsoon rainfall as independent variable. Column (2) also controls for state- by NREGA implementation phase linear time trends. Columns (3) and (4) are estimated on a balanced district level annual panel. Column (3) is a linear probability model using a dummy variable as dependent variable indicating whether a district experienced any conflict events in a given year. Column (4) estimates a Poisson regression with the dependent variable being the number of conflict events per district and year. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. For columns (1)-(3) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Robustness to Adding Controls: Moderating Effect of NREGA on Conflict

	Incidence			Intensity			
	(1) Controls	(2) NREGA x FE	(3) State x Year FE	(4) Prev Conflict	(5) Controls	(6) NREGA x FE	(7) State x Year FE
log(Monsoon)	-0.019	-0.030**	-0.044**	-1.360***	-0.800***	-0.930***	-0.455
	(0.019)	(0.014)	(0.018)	(0.305)	(0.240)	(0.295)	(0.362)
NREGA x log(Monsoon)	-0.008	0.011	0.008	0.975**	0.951***	1.133***	0.250
	(0.014)	(0.028)	(0.013)	(0.415)	(0.327)	(0.385)	(0.340)
Observations	7059	7059	7059	1794	2760	2217	2580
Number of Districts	543	543	543	144	222	222	222

Notes: Data is on a balanced panel of conflict events from 2000-2012. Column (1)-(3) are linear probability models using a dummy variable as dependent variable indicating whether a district experienced any conflict events in a given year. Column (4) - (7) estimates a Poisson regression with the dependent variable being the number of conflict events per district and year. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. Column (1) and (5) include controls interacted with year fixed effects. The district characteristics are: terrain ruggedness, elevation, rural population share, tribal population share, scheduled caste share, illiteracy rates, household size, share of population younger than 6 years, population growth rate from census 1991 to 2001, gender gap, share of villages in district with primary school, share of villages in district with mud road approach, share of households in district that live in permanent housing, share of villages in district with primary health care facilities, share of villages with electric power, share of villages with a bus stop and the share of villages with a postal office. Columns (2) and (6) interact the NREGA treatment dummy with district fixed effects. Columns (3) and (7) control for state by year fixed effects. In column (4) I constrain the analysis to the districts that experienced conflict prior to NREGA. For columns (1)-(3) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Robustness to Weather Measures: Moderating Effect of NREGA on Conflict

	Incidence				Intensity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Normalized Monsoon	-0.013**				-0.225***			
	(0.006)				(0.043)			
NREGA x Normalized Monsoon	0.006				0.157***			
	(0.005)				(0.057)			
log(GPCC Rain)		-0.031**				-1.437***		
		(0.015)				(0.277)		
NREGA x log(GPCC Rain)		0.040***				1.695***		
		(0.012)				(0.352)		
Fitted log(Monsoon)			-0.042*				-1.208***	
			(0.023)				(0.401)	
NREGA x Fitted log(Monsoon)			0.047***				0.981**	
			(0.014)				(0.441)	
NDVI				-0.287*				-2.918
				(0.172)				(2.987)
NREGA x NDVI				0.154***				2.222*
				(0.042)				(1.176)
Observations	7059	7059	7059	6516	2760	2760	2760	2536
Number of Districts	543	543	543	543	222	222	222	219

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects. Notes: Data is a balanced panel of conflict events from 2000-2012. All weather measures are lagged by one year. Column (1)-(4) are linear probability models using a dummy variable as dependent variable indicating whether a district experienced any conflict event in a given year. Column (5) - (8) estimates a Poisson regression with the dependent variable being the number of conflict events per district and year. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. Columns (1) and (4) present results where Monsoon rainfall is normalized by its standard deviation. Columns (2) and (5) use the GPCC rainfall data as alternative rainfall data source. Columns (3) and (6) instrument the TRMM rainfall data with the GPCC data to remove measurement error. Columns (4) and (8) use the Modis Vegetation index as measure of photosynthetic activity available from 2000-2011. For columns (1)-(4) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions in columns (5) - (8) present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Robustness to Treatment Timing: Moderating Effect of NREGA on Conflict

	Incidence				Intensity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Monsoon)	-0.048***	-0.050***	-0.050***		-1.166***	-1.043***	-1.002***	
	(0.018)	(0.018)	(0.018)		(0.310)	(0.334)	(0.374)	
NREGA Placebo 1 x log(Monsoon)	0.029*				0.627*			
	(0.015)				(0.370)			
NREGA Placebo 2 x log(Monsoon)		0.030**				0.409		
		(0.014)				(0.378)		
NREGA Placebo 3 x log(Monsoon)			0.026*				0.310	
			(0.014)				(0.385)	
log(Outside Monsoon)				-0.002				-0.216
				(0.008)				(0.264)
NREGA x log(Outside Monsoon)				-0.001				0.162
				(0.004)				(0.316)
Observations	7059	7059	7059	7059	2760	2760	2760	2760
Number of Districts	543	543	543	543	222	222	222	222

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects. Column (1)-(4) are linear probability models using a dummy variable as dependent variable indicating whether a district experienced any conflict event in a given year. Column (5) - (8) estimates a Poisson regression with the dependent variable being the number of conflict events per district and year. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. Placebo 1 -3 move the NREGA treatment indicator 1, 2, 3 years ahead of time, respectively. For columns (1)-(4) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Extended Results: Effects on Overall Crime

	(1) All	(2) Violent	(3) Property	(4) Public Order	(5) Women
log(Monsoon)	0.009	-0.061**	0.033	-0.118***	0.059*
	(0.021)	(0.031)	(0.029)	(0.042)	(0.032)
NREGA x log(Monsoon)	0.028**	0.072***	-0.016	0.165***	0.039
	(0.014)	(0.019)	(0.020)	(0.044)	(0.029)
Observations	5356	5356	5356	5356	5356
Number of Districts	537	537	537	537	537
F-test	2.64	.09	.2	.92	6.15
p-value	.1	.76	.65	.34	.01

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation. The dependent variable is the log of the number of reported crime incidents in the category given in the column head per district and year from 2002-2012. Standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Explaining the NREGA Effect: Monsoon Rainfall and Targets of Violence

	Incidence			Intensity		
	(1) Civilian	(2) Security	(3) Terrorist	(4) Civilian	(5) Securitry	(6) Terrorist
log(Monsoon)	-0.061***	-0.021**	-0.019	-1.532***	-0.950**	-1.196***
	(0.017)	(0.009)	(0.012)	(0.309)	(0.392)	(0.346)
NREGA x log(Monsoon)	0.044***	0.026***	0.022*	1.352***	1.317***	0.622
	(0.013)	(0.010)	(0.012)	(0.402)	(0.500)	(0.407)
Observations	7059	7059	7059	2490	1673	1893
Number of Districts	543	543	543	203	137	153

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects. Columns (1)-(3) are linear probability models using a dummy variable as dependent variable indicating whether a district experienced any conflict events in a given year. Columns (4)-(6) estimate Poisson regressions with the dependent variable being the number of conflict events per district and year. Conflict events are categorized into whether the subject of a conflict event was a civilian, security force or terrorists. Note that conditional fixed effect poisson models drop districts which do not have any variation in the dependent variable. For columns (1)-(3) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Explaining the NREGA Effect: Monsoon Rainfall and NREGA Participation

	Costs	Participation			Heterogeneity	
	(1) Projects	(2) Overall	(3) Extensive	(4) Intensive	(5) Phase	(6) LWE
log(Monsoon)	-0.257**	-0.212***	-0.055***	-0.114**	-0.078***	-0.030**
	(0.101)	(0.075)	(0.014)	(0.056)	(0.024)	(0.014)
Phase 2 x log(Monsoon)					0.046*	
					(0.025)	
Phase 3 x log(Monsoon)					0.038	
					(0.025)	
LWE Affected x log(Monsoon)						-0.056**
						(0.024)
Observations	2872	3060	3060	3060	3060	3060
Number of Districts	504	537	537	537	537	537

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation. Column (1) studies the log of total cost of achtive NREGA projects, columns (2)-(4) study participation as log of total person days employed in column (2), share of households in a district in column (3) and the log of number of days employed per household in column (4). Columns (5) and (6) study heterogeneity in extensive margin participation. Standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10: Insurance Value of NREGA: Monsoon Rainfall, Output Losses and NREGA Expenditures

	Output Value/Capita	NREGA Expenditure/Capita	
	(1) OLS	(2) OLS	(3) IV
log(Monsoon$_t$)	541.335***		
	(160.249)		
log(Monsoon$_{t-1}$)		-101.219***	
		(32.984)	
Output Value/ Capita$_{t-1}$			-0.308***
			(0.103)
First Stage			12.2
Observations	4086	3059	1465
Number of Districts	438	537	410

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Column (1) relates Monsoon rainfall with agricultural output per capita. Column (2) studies lagged Monsoon rainfall and its effect on levels of NREGA expenditure in a district per capita. Column (3) is an instrumental variables exercise, instrumenting lagged agricultural output value per capita with lagged Monsoon rainfall. Standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Extended Results and Further Robustness

A.1 Other Explanations

This section rules out a range of alternative explanations or policies that could explain why the Monsoon rainfall and conflict relationship has become weaker. Most notably is the Pradhan Mantri Gram Sadak Yojana Rural Road Construction Scheme (PMGSY) that was implemented around the same time as NREGA was devised and introduced. Other confounders are the Integrated Action Plan, which channels additional funds into left-wing extremist affected districts. I also address a concern that large mineral sectors are driving the observed moderation. Lastly, I address issues concerning a Military Operation that has been underway since early 2011.

Pradhan Mantri Gram Sadak Yojana Rural Road Construction Scheme A concern with the analysis is that the Indian government has put forth many other development programs, whose implementation may affect the relationship between Monsoon and conflict at the same time and may be correlated with the roll-out of NREGA. In this case, the results would falsely attribute the observed inward rotation of the Monsoon-rainfall andconflict relationship to the employment guarantee scheme. The most prominent developmental scheme that was implemented around the same time is the Pradhan Mantri Gram Sadak Yojana (PMGSY). This scheme was introduced in 2000 and aims to provide improved road access for rural households. The scheme in particular aimed to provide roads to all villages with at least 1000 inhabitants by 2003, with a population of 500 and more by 2007 and had special provisions for tiny villages with at least 250 inhabitants for the hill states, tribal areas and desert areas. These were to be connected by 2007. As early NREGA districts are among the poorest and least urbanised, they are more likely to have received treatment through the PMGSY as well, which could partly explain my reduced form findings.

The crucial role that transport infrastructure may have in mitigating adverse weather shocks has been highlighted in Burgess and Donaldson (2010). Aggarwal (2014) evaluates the impact of the PMGSY using a difference-in-difference design and finds that the scheme increased incomes by increasing the potential market size for locally produced agricultural commodities; in addition, there is less price dispersion across market centers. I use her data to see whether the PMGSY moderates the relationship between Monsoon rainfall and conflict. I construct two variables: first, the share of all unconnected habitats connected in a year and second, the cumulative share of habitats among the unconnected habitats that received road access by the end of each year. The former measure may pick up direct effects from road construction on violence, while the latter variable, in its interaction with rainfall, could pick up the more persistent effects of this scheme by connecting previously unconnected villages.

The empirical design is identical to the main analysis, except that I now add these controls and interaction terms to the main specification. The results are presented in

Table A5. Column (1) and (2) study violence intensity, while column (3) and (4) look at incidence. Panel A presents the results for contemporary road construction, while Panel B looks at cumulative connectivity. Columns (1) and (3) look at the rural connectivity and its interaction with rainfall by itself, while column (3) and (4) are a type of horse race. In neither specifications do the road construction interactions with rainfall achieve predictive power. This renders me confident that my results genuinely reflect the effect of the workfare scheme on the dynamics of conflict.

Integrated Action Plan A second important policy aimed to tackle the Naxalite conflict is the Integrated Action Plan (henceforth, IAP). The plan was presented in 2010 and provides special funding for for districts that are considered to be severely affected by left-wing extremism. Originally it was designed for 33 districts, but since then, it expanded to provide additional funding for 82 districts. The money is to be spend on projects such as roads and other public infrastructure to improve rural livelihoods; some projects are specifically aimed to improving the way NREGA is made accessible in these districts: some IAP funding may be used to complement NREGA projects. Another margin through which the IAP may have a distinct level effect on conflict is provided as money may be used to reinforce police stations to expand the states' presence in rural areas.

Investment in infrastructure funded by the IAP could moderate the rainfall dependence of income and thus, on conflict. I don't think that the IAP would have the effects described in this paper, as its implementation would have to correlate meaningfully with lagged Monsoon rainfall. Since the grants are block grants, this is unlikely to be the case. Nevertheless I study this and the results are presented in Table A6. In any case, there are three simple things I can do to rule out effects of the IAP driving my results. Firstly, I can drop the 33 districts which received the scheme from 2010 onwards.[49] The results from this is presented in columns (2) and (5). The interaction term becomes smaller and size and statistical significance, especially for the conflict intensity regressions. This is not implausible as the districts that receive the IAP are ones with most variation in conflict. In second exercise, I can restrict the analysis to the period from 2000 - 2010. Again, the estimated coefficient on the post NREGA period become weaker, but the core result is still there. In the last exercise I study IAP fund expenditures, which measures utilisation of the disbursal amounts. Column (1) indicates that IAP expenditures are not correlated with lagged Monsoon rainfall. Column (3) and (6) study the effect of IAP expenditures on conflict. There appears to be a positive relationship between the two. The estimated coefficient on the NREGA interaction remains the same, thus rendering the core result robust.

[49] The districts translate into 30 districts according to the 2001 Indian census district definitions, they are: Aurangabad (Bihar), Arwal, Balaghat, Bastar, Bokaro, Chatra, Dantewada, Deogarh, Gadchiroli, Gajapati, Gaya, Garhwa, Gondiya, Gumla, Hazaribagh, Jamui, Jehanabad, Khammam, Lohardaga, Midnapore, Nabarangpur, Palamu, Pashchim singhbhum, Purba singhbhum, Rajnandgaon, Rayagada, Rohtas, Sambalpur, Sonbhadra, Surguja, Malkangiri.

Operation Green Hunt A major military operation to tackle Naxalite violence has been underway since late November 2009. The operation involves the deployment of Central Armed Reserve Police force to aide state governments tackle Naxalite threat. If deployment of troops is correlated with lagged Monsoon season rainfall, this could explain some of the observed patterns. It is not clear in which direction the effect should be. If military deployment was correlated with lagged rainfall, increased military deployment following an adverse shock could either lead to a conflict escalation or a reduction in conflict. Unfortunately, data on military deployment is not available. As with the integrated action plan, I can limit the analysis to the period before 2010 or by removing a set of districts that likely, were the primary target for a military operation. The second main concern is the relationship between the local prevalence of mineral resources. Mineral resources are a natural stabiliser to district level income, as the resource revenues are less likely to depend strongly on Monsoon rainfall. As the period around 2007 saw a major commodity boom, this could have boosted mineral resource revenues. While it is not clear that this was correlated with Monsoon rainfall in a systematic manner, it is still an important to assess the relevance of this channel as it relates this paper back to the existing literature.

Mineral Resources Another concern is that the NREGA interactions may be picking up moderation of rainfall shocks due to a sectoral shift away from agriculture to the mining sector, which is less affected by rainfall variation. Vanden Eynde (2011) shows that districts with a large mining sector see a smaller elasticity between rainfall and conflict. If the introduction of NREGA is correlated with a sectoral shift towards the mineral resource sector, the NREGA interactions could be picking up this effect. This is not entirely implausible as the mid 2000s saw a commodity price boom which could have induced a lot more investment in the mining sector. In order to control for this I construct a share of a district's income that is due to the mining sector.[50]

Again, the specifications I present are very similar, adding a simple interaction with the mining sector share in district domestic product interacted with the Monsoon season rainfall. The results are presented in Table A7. Column (1) presents the results on violence intensity without the NREGA interactions. It becomes evident that districts with a larger share of the mining sector experience a weaker relationship between violence and Monsoon rain. This maps into the findings of Vanden Eynde (2011). Once including the NREGA interaction, the coefficient on the Mining sector interaction becomes insignificant with a p-value of 12%. More importantly, the NREGA interaction remains strongly significant. This suggests that the NREGA effect seems not to be picking up a moderation in the Monsoon shock and conflict relationship due to the presence of a large mineral resource sector.

[50] I use district domestic product data for years between 1998 to 2005 available from `http://planningcommission.nic.in/plans/stateplan/index.php?state=ssphdbody.htm`, accessed on 21.06.2014. The district domestic product construction is discussed in detail in Katyal et al. (2001).

The next section provides tentative results of a level effect of NREGA on conflict levels.

A.2 Level Effect of NREGA

The preceding results suggested that NREGA does have a moderating effect on the cyclical nature of violence, in particular, the violence targeted against civilians. However, the existing literature evaluating the economic impacts of NREGA also indicate strong increases in wage levels.[51] An increase in wage levels can be seen as an increase in the returns to labour in both, good- and bad states of the world. This does have an independent level effect on conflict. It is challenging to identify a level effect due to the endogeneity of the roll-out. Nevertheless, in this section I provide an estimate of the level effect of NREGA. I estimate specifications with less demanding time-fixed effects that vary by region. This ensures that the coefficient on the NREGA treatment dummy is not collinear with the time effects and can thus, be interpreted. The specification I estimate is as follows:

$$\mathbb{E}(A_{dprt}) = \delta_d \exp\left(b_{rt} + \alpha T_{dprt} + \eta R_{dpr,t-1} + \gamma T_{dprt} \times R_{dpr,t-1} + \epsilon_{dprt}\right) \quad (8)$$

where b_{rt} are now region by time fixed effects, rather than region by phase and time fixed effects. This set of fixed effects allows the estimation of the parameter α, which can be interpreted as the level effect of NREGA if we are willing to assume that the roll-out of NREGA was exogenuous. This is not a classical difference in difference estimator with one set of treated and one set of untreated locations since eventually, all districts receive NREGA. The coefficient α is estimated off the time variation due to the sequential roll out of NREGA. This implies that the NREGA treatment indicator is estimated soley from the variation for the years in which some districts had already received NREGA relative to other districts that did not have NREGA yet; that is, the coefficient solely lives off the variation in differences in conflict across phases for the years 2006 and 2007. When adding interaction terms with Monsoon rainfall and NREGA, the interaction term becomes, in addition, a heterogenous effect for the level effect of NREGA in these two years. In order to get the average treatment effect I demean Monsoon rainfall variable for these regressions. I estimate three versions of the above specification. First, imposing the constraint that $\eta = \gamma$. In this case, I force the effect of rainfall to be the same before and after the introduction of NREGA. I also estimate a specification with the constraint $\eta = \gamma = 0$, which effectively means not controlling for rainfall. The key question is how this will affect the estimated coefficient $\hat{\alpha}$. In both cases, the coefficient $\hat{\alpha}$ should overstate the effect of NREGA in absolute value.

The results are presented in table A8. The first column presents the constrained regression where I do not control for rainfall. The level effect coefficient is negative and statistically significant. This coefficient is a mixture of the level effect and the implied

[51] See Zimmermann (2012), Berg et al. (2012), Imbert and Papp (2014) and Azam (2011).

effect due to a reduced rainfall and conflict elasticity. In the second column, I control for rainfall, which renders the coefficient slightly larger in absolute value. The third column is the unconstrained coefficient, allowing the functional relationship between rainfall and conflict to change with the introduction of NREGA. The interesting observation is that the coefficient on the level effect goes down and is estimated relatively imprecisely, moving from a p-value close to 0.001 to p-value of 0.45. This suggests that the dynamic effect of NREGA, operating by mitigating income shocks, is being partially captured in estimates of $\hat{\alpha}$, when one does not explicitly control for this important economic channel through which NREGA operates. When comparing column (2) and column (3), this suggests that at least 1/3 of the estimated reduction in violence levels is due to the reduced rainfall dependence of conflict.

This paper provides evidence that NREGA functions as insurance. This suggests that the correct way to evaluate NREGA is through its dynamic effect through program participation. Nevertheless recently, Dasgupta et al. (2014) and Khanna and Zimmermann (2013) separately estimated level effects of the introduction of NREGA. They arrive at different conclusions. Khanna and Zimmermann (2013) use a regression-discontinuity design relying on a reverse engineered NREGA roll-out algorithm to identify districts that were close to the cutoff of being assigned into either an earlier, or a later phase. They argue that this provides a good counterfactual for a fuzzy regression discontinuity design and estimate the effect of the NREGA treatment. They find that NREGA increased conflict levels in the short-run. Dasgupta et al. (2014) use a difference in difference estimator as I discussed above. This design lives off variation in differences in conflict arising due to the gradual roll-out. I provide some evidence of level effects in this paper, estimating a similar difference-in-difference specification as in Dasgupta et al. (2014). The results are presented in Table A9. The first column presents the basic level effect estimate of contemporaneous treatment. The second column adds lagged effects of the NREGA treatment indicator, suggesting that the first lag is highly significant. The point estimate suggest that the introduction of NREGA reduced levels of violence by between 30% to 50% for average Monsoon rainfalls. Columns (4)-(9) explore the heterogeneity of the estimated effect by interacting the treatment indicator with a set of district-characteristics. The district characteristics are demeaned for ease of interpretation of the marginal effects. The results suggest that the level effect is weaker for districts with a high scheduled tribe share, but stronger for districts with higher scheduled caste share. Indicative is the coefficient on average household size. This suggests that the level effect is significantly weaker for districts with a larger average household size. Since the NREGA program provides an allowance for 100 days of work *per household*, larger households are disadvantaged in that respect. Column (8) interacts the treatment indicator with the the mean level of agricultural output per capita before 2005 expressed in INR 1000. The coefficient is negative and significant, suggesting that richer districts saw a stronger drop in conflict. While the results on the dynamics of conflict do not square with Khanna and Zimmermann (2013), the estimated level effects do stand at odds with the ones estimated in their paper but map well into the findings of Dasgupta et al. (2014). That being said, as NREGA appears to provide insurance, the correct way to evaluate the impact NREGA should be through its dynamic effect on conflict, rather than through its static level effect.

A.3 NREGA Road Construction and Conflict

NREGA aims to "create durable assets which have potential to generate additional employment in the years to come in rural areas."[52] While the analysis of the agricultural production function did not suggest a dramatic change in the relationship between Monsoon rainfall and agricultural output in the short run, asset construction under NREGA could still affect the dynamics of conflict. There is anecdotal evidence suggesting that Naxalites oppose road construction under the scheme (see Banerjee and Saha (2010)). The anecdotal accounts suggest that this is for fear that roads could provide easier access for police and military. There is anecdotal evidence suggesting that Naxalites have taken road construction contractors hostage or killed them, suggesting that road construction could drive conflict.[53] There are two ways that road-construction could affect the results here. First, road construction may itself be correlated with lagged Monsoon rainfall and through that affect the dynamics of conflict in a way that is correlated with Monsoon rainfall. There could also be an independent effect from road construction that affects conflict levels. This section shows two things. I show that road construction is correlated with lagged Monsoon rainfall; however, this relationship is not present for districts for which the moderation in the rainfall and conflict relationship is strongest. There appears to be a distinct effect of road construction on conflict that is not related to Monsoon season rainfall. Districts in which the share of overall NREGA funds allocated to road construction in the years since NREGA was introduced is higher, experience more conflict in recent years. This effect is a mere correlation due to the endogeneity of NREGA road construction. It is however, indicative for further research.

Monsoon Rainfall and Road Construction If road construction itself was correlated with lagged Monsoon rainfall, this could explain the finding of the inward rotation of the relationship between Monsoon rainfall and conflict. The argument is quite simple. Before NREGA, good Monsoon rains would reduce conflict. With NREGA available, strong Monsoon rainfalls may be associated with increasing road construction to repair mud roads that have been damaged due to the Monsoon. This may lead to more conflict, reversing the previously existing relationship. This is a genuine concern and is studied in this section.

[52] See http://www.nrega.nic.in, accessed 12.02.2014.

[53] See for example http://www.satp.org/satporgtp/detailsmall_news.asp?date1=7/16/2011&id=5, accessed 11.10.2014.

A brief look at summary statistics is already quite telling. Studying districts that have been categorized as being under left wing extremist influence or have seen some conflict for the period prior to NREGA suggests that these districts see a significantly lower share of NREGA expenditure going to road construction. The total expenditure weighted share in 2010 is around 31.3%, while it is 36.3% for the other districts. This is an important insight, as it suggests that the types of assets created under NREGA may reflect local preferences.

Table A10 presents results studying how the share of overall NREGA expenditures in road construction (columns 1-3) or for land development (column 4-6) in a financial year respond to Monsoon rainfall in the previous season. The relationship suggests that lagged Monsoon rainfall predicts an increasing share of road construction (column 1), but not for projects that can be classified for land development (column 4). This gives rise to the genuine concern that rainfall may drive road construction which, in turn, is driving conflict as suggested. I rule out this explanation by studying the heterogeneity across NREGA implementation phase, by whether districts are classified as being under left-wing extremist influence and by studying whether it is excessive Monsoon rains that drive this effect.

Columns (2) and (4) study this relationship for districts by implementation phase: it appears that the positive Monsoon rainfall and road construction relationship is driven by districts that receive NREGA in early phases. Columns (3) and (6) study the responsiveness of NREGA asset construction for districts classified under left-wing extremist influence. Neither expenditures for land-development (column 6) nor road construction (column 3) meaningfully correlate with Monsoon rainfall for these districts. This is reassuring given that the moderating effect of NREGA on the Monsoon conflict relationship is coming mostly from these districts.

A non-parametric analysis further suggests that it is positive rainfall that correlates with road construction for phase 1 districts (see Figure A1). This is an important insight as the non-parametric analysis of the Monsoon rainfall and conflict (or crime) relationship indicated that NREGA's moderating effect on that relationship is due its impact on below normal Monsoon rainfall. This renders me confident that, while road construction may have an independent effect on conflict, this effect is not confounding the moderation in the Monsoon rainfall and conflict relationship studied in this paper.

This is also conceptually reasonable: road construction due to excessive Monsoon may simply repair and replace already existing mud roads, which are most prevalent in districts that received NREGA in the first phase (see Table 1). This is qualitatively a lot different from new roads being constructed. Places, in which a lot of roads are constructed, may experience a change in the conflict dynamic, as it is new roads that improve access to remote places for the military and not the improvement or repairing of existing roads. That is to say: places that receive a lot of road construction through NREGA independent of Monsoon rainfall may experience a change in the conflict dynamics that is, however, unrelated to the moderation in the Monsoon and conflict relationship studied here. This is highlighted in the next paragraph.

Independent Effect of New Road Construction In order to study the direct effect of road construction on conflict that is unrelated to Monsoon rainfall, I construct a measure of road construction intensity as the share of all funds devoted to road construction activity for all the post NREGA years. This overall measure may reflect local preferences for different development projects that is, due to the averaging, independent of Monsoon rainfall. Let this measure be denoted as ρ_d. I estimate the event study analysis interacting the time to treatment with the measure ρ_d and plot out the coefficients. The coefficients are estimated off the variation in NREGA road construction intensity across districts. The results are presented in Figure A2. The pattern that emerges suggests that NREGA road construction intensity is correlated with higher incidence and intensity of conflict following the introduction of NREGA.

The second column indicates that both conflict incidence and intensity is positively correlated in districts that see a large share of their NREGA funds devoted to road construction. This is indicative of there being a distinct channel through which road construction leads to more conflict. However, controlling for this channel leaves the moderating effect of the program on the relationship between Monsoon rainfall and conflict intact. This is presented in the left column. The moderation in the Monsoon rainfall and conflict relationship is still present and significant.

Figures and Tables for Robustness Appendix

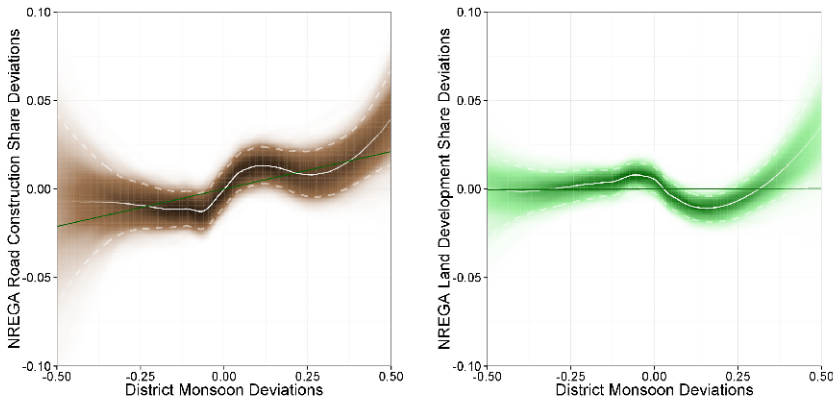

Figure A1: NREGA Infrastructure Expenditure Shares and Lagged Monsoon Rainfall for Phase 1 Districts

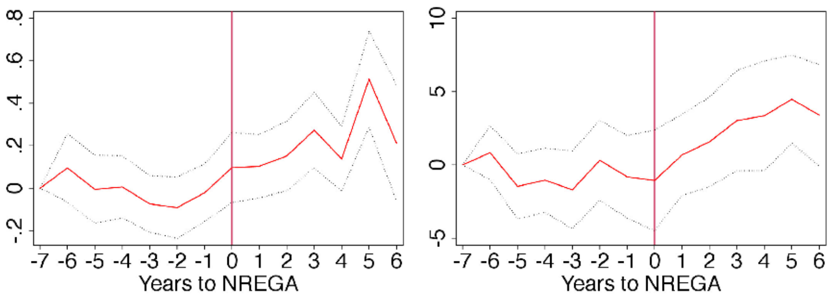

Figure A2: Effect of NREGA Cumulative Road Construction Expenditure Share on Conflict over Time.

Table A1: Socio-Economic Characteristics of Districts with Naxalite Presence

	LWE Affected	Other Districts
Number of Districts	206	365
Panel A: Demographic		
Rural Population	77.17	69.52
Tribal Population	9.47	7.49
Scheduled Caste	16.55	15.75
Illiterate	46.07	44.56
Population Age < 6	24.39	24.81
Permanent House	46.69	55.81
Panel B: Infrastructure		
Primary School	74.27	81.57
Mud Road	72.11	59.63
Primary Health Care	29.72	33.19
Electricity	65.41	84.86
Bus Stop	29.66	40.35
Post Office	34.03	46.83

Notes: Statistics derived from the 2001 Census for India. Panel A presents demographic indicators as shares of the overall population. Panel B presents Infrastructure indicators derived from the share of villages that have access to a particular type of infrastructure.

Table A2: Before the Introduction of NREGA: Robustness of Relationship between Weather Variables and Agricultural Output

	log(Output Value/Capita)			log(Grain Value/Capita)		
	(1) Outside Monsoon	(2) Temperature	(3) Controls	(4) Outside Monsoon	(5) Temperature	(6) Controls
log(Monsoon)	0.364***	0.364***	0.424***	0.369***	0.369***	0.335***
	(0.086)	(0.087)	(0.080)	(0.076)	(0.076)	(0.073)
log(Outside Monsoon)	0.122**	0.124**	0.129***	0.114***	0.115***	0.090**
	(0.051)	(0.051)	(0.046)	(0.043)	(0.043)	(0.042)
Temperature		-0.063	-0.030		-0.028	-0.053*
		(0.051)	(0.031)		(0.040)	(0.030)
Observations	3239	3239	3239	3196	3196	3196
Number of Districts	471	471	471	464	464	464

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Temperature measures the average temperature during the Monsoon months. Columns (1)-(3) study agricultural output value per capita, while columns (4) - (6) study the value of grain production encompassing ragi, rice, wheat, bajra, jowar, maize, pules and barley. Columns (3) and (6) add a set of district characteristics interacted with a set of year fixed effects. The district characteristics are: terrain ruggedness, elevation, rural population share, tribal population share, scheduled caste share, illiteracy rates, household size, share of population younger than 6 years, population growth rate from census 1991 to 2001, gender gap, share of villages in district with primary school, share of villages in district with mud road approach, share of households in district that live in permanent housing, share of villages in district with primary health care facilities, share of villages with electric power, share of villages with a bus stop and the share of villages with a postal office. Standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: Before the Introduction of NREGA: Robustness of Relationship between Weather Variables and Agricultural Wages

	log(Annual Wage)			log(Seasonal Wage)	
	(1) Outside Monsoon	(2) Temperature	(3) Controls	(4) Harvesting	(5) Planting
log(Monsoon)	0.058***	0.058***	0.057***	0.046*	0.010
	(0.019)	(0.018)	(0.020)	(0.028)	(0.017)
log(Outside Monsoon)	0.006	0.006	0.011		
	(0.011)	(0.011)	(0.015)		
Temperature		0.001	0.021		
		(0.015)	(0.013)		
State by NREGA Phase Trend	Yes	Yes	No	Yes	Yes
Observations	1419	1419	1419	1387	1195
Number of Districts	314	314	314	318	260

Notes: All regressions include region-phase-time fixed effects and district fixed effects. State by NREGA-Phase Trend are linear trends at the State by NREGA implementation phase level. Temperature measures the average temperature during the Monsoon months. Data is an unbalanced district level panel of annual agricultural wages in India. Columns (1)-(3) study agricultural wages, while columns (4) and (5) study wages at the planting stage compared to wages at harvesting stage towards the end of the year. Column (3) adds a set of district characteristics interacted with a set of year fixed effects. The district characteristics are: terrain ruggedness, elevation, rural population share, tribal population share, scheduled caste share, illiteracy rates, household size, share of population younger than 6 years, population growth rate from census 1991 to 2001, gender gap, share of villages in district with primary school, share of villages in district with mud road approach, share of households in district that live in permanent housing, share of villages in district with primary health care facilities, share of villages with electric power, share of villages with a bus stop and the share of villages with a postal office. Standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: Before the Introduction of NREGA: Robustness of Relationship between Monsoon Rainfall and Conflict

	Robustness to Choice of Empirical Model			Controls and Interactions		
	(1) Poisson-IV	(2) Neg Bin	(3) OLS	(4) Weather	(5) Controls	(6) Interactions
log(Fitted Output Value/Capita$_{t-1}$)	-2.362**					
	(0.986)					
log(Monsoon$_{t-1}$)		-0.830***	-0.313**	-0.960***	-0.811**	-1.192***
		(0.230)	(0.147)	(0.346)	(0.328)	(0.422)
log(Outside Monsoon$_{t-1}$)				-0.384*		
				(0.205)		
Temperature$_{t-1}$				0.504		
				(0.314)		
log(Monsoon$_t$)				-0.056		
				(0.257)		
Temperature$_t$				0.096		
				(0.310)		
District Controls	No	No	No	No	Yes	No
Monsoon Rain Interactions	No	No	No	No	No	Yes
Observations	646	932	3843	932	932	932
Number of Districts	117	144	543	144	144	144

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Column (1) presents the results of an IV regression, instrumenting lagged agricultural output per capita with lagged Monsoon rainfall. Column (2) is a negative binomial, while column (3) presents OLS results. Column (4) includes temperature during the Monsoon season as well as contemporaneous weather. Column (5) interacts a set of district controls with a set of year fixed effects, while column (6) interacts Monsoon rainfall with the demeaned district characteristics. The district characteristics are: terrain ruggedness, elevation, rural population share, tribal population share, scheduled caste share, illiteracy rates, household size, share of population younger than 6 years, population growth rate from census 1991 to 2001, gender gap, share of villages in district with primary school, share of villages in district with mud road approach, share of households in district that live in permanent housing, share of villages in district with primary health care facilities, share of villages with electric power, share of villages with a bus stop and the share of villages with a postal office. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: Alternative Mechanism: Rural Connectivity and Moderation of Rainfall and Conflict Relationship

	Incidence		Intensity	
	(1)	(2)	(3)	(4)
Panel A: Road Construction				
log(Monsoon)	-0.040**	-0.059***	-0.712***	-1.418***
	(0.020)	(0.022)	(0.212)	(0.358)
NREGA x log(Monsoon)		0.047***		1.093***
		(0.017)		(0.394)
Roads	-0.387	-0.622	5.127	4.153
	(0.571)	(0.571)	(5.249)	(5.203)
Roads x log(Monsoon)	0.073	0.108	-0.647	-0.521
	(0.092)	(0.091)	(0.753)	(0.756)
Panel B: Cumulative Road Construction				
log(Monsoon)	-0.049**	-0.055**	-0.927***	-1.325***
	(0.023)	(0.024)	(0.311)	(0.347)
NREGA x log(Monsoon)		0.044***		1.415***
		(0.016)		(0.354)
Cumulative Roads	-0.248	0.005	-7.788*	4.778
	(0.248)	(0.226)	(4.647)	(4.011)
Cumulative Roads x log(Monsoon)	0.040	0.002	0.917	-0.880
	(0.037)	(0.034)	(0.661)	(0.586)
Observations	5775	5775	2230	2230
Number of Districts	525	525	210	210
Estimation	OLS	OLS	Poisson	Poisson

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation. The dependent variable is the number of violent incidences per quarter in columns (1) and (2) and an indicator whether there was any violent incidence in columns (3) and (4). Panel A studies the effect of contemporaneous road construction on violence, while Panel B studies the impact of rainfall through the overall share of unconnected habitats that became connected up to 2012. Standard errors are clustered at district level in column (1) and (2), while in column (3) and (4) they are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: Alternative Mechanism: Integrated Action Plan Disbursals and the Moderation of Monsoon Rainfall and Conflict Relationship

	IAP	Incidence			Intensity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(Monsoon)	0.407	-0.045***	-0.044***	-0.048***	-1.257***	-1.170***	-1.352***
	(0.426)	(0.017)	(0.015)	(0.018)	(0.279)	(0.271)	(0.286)
NREGA x log(Monsoon)		0.037***	0.046***	0.042***	0.642*	0.880**	1.008***
		(0.014)	(0.016)	(0.014)	(0.366)	(0.424)	(0.384)
IAP Expenditure				0.011			0.084***
				(0.010)			(0.029)
Observations	184	6669	5973	7059	2354	2152	2760
Number of Districts	72	513	543	543	192	205	222
Estimation	OLS	OLS	OLS	OLS	Poisson	Poisson	Poisson

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realization. Column (1) studies IAP expenditure as a function of lagged Monsoon rain. The dependent variable in columns (2)-(4) is an indicator whether there was any conflict event in a district and year, while it is the number of conflict events per year in columns (5)-(7). Columns (2) and (5) remove the 33 districts that received the IAP originally. Columns (3) and (6) restrict the analysis to the period 2000-2010. Columns (4) and (7) control for IAP expenditure. Standard errors in column (1)-(4) are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. Errors in columns (3) and (4) are clustered at the district level, with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: Alternative Mechanism: Mining Sector Share, Commodity Boom and Moderation of Rainfall and Conflict Relationship

	Incidence		Intensity	
	(1)	(2)	(3)	(4)
log(Monsoon)	-0.042**	-0.054***	-0.947***	-1.797***
	(0.018)	(0.018)	(0.199)	(0.283)
NREGA x log(Monsoon)		0.049***		1.369***
		(0.015)		(0.389)
Mining Sector Share	0.011	0.012	4.554*	3.820*
x log(Monsoon)	(0.315)	(0.313)	(2.348)	(2.256)
Observations	6552	6552	2504	2504
Number of Districts	504	504	204	204
Estimation	OLS	OLS	Poisson	Poisson

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation. The dependent variable is an indicator whether there was any conflict event in columns (1) and (2) and the number of violent incidences per year in columns (3) and (4). Mining Sector Share is the share of the districts domestic product that is generated in the Mining sector based on data between 1998 and 2005. Standard errors in column (1) and (2) are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. Errors in columns (3) and (4) are clustered at the district level, with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A8: Dynamic Versus Direct Level Effect of NREGA

	(1) $\eta = \gamma = 0$	(2) $\eta = \gamma$	(3) Unconstrained
NREGA	-0.437***	-0.498***	-0.300*
	(0.165)	(0.178)	(0.170)
log(Monsoon)		-0.859***	-1.758***
		(0.237)	(0.356)
NREGA x log(Monsoon)			1.457***
			(0.369)
Observations	2886	2886	2886
Number of Districts	222	222	222

Notes: All regressions include region-by-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation; Monsoon rainfall is demeaned for ease of interpretation of the interaction terms. All regressions are estimated using Poisson models with the dependent variable being the number of conflict events per district and year. The first column does not control for Monsoon rainfall, while the second column constraints the rainfall coefficient to be the same before, and after the introduction of NREGA. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A9: Level Effect of NREGA

	Level Effect Estimates			Heterogeneity of Level Effect				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NREGA	-0.300*	-0.225	-0.533***	-0.724***	-0.297*	-0.263	-0.695***	-0.930***
	(0.170)	(0.150)	(0.197)	(0.198)	(0.173)	(0.171)	(0.234)	(0.232)
$\text{NREGA}_{t\text{-}1}$		-0.470**						
		(0.201)						
Heterogeneity: NREGA $\times$								
Scheduled Tribe			1.051***					0.540
			(0.400)					(0.448)
Scheduled Caste				-5.974***				-2.651
				(1.539)				(2.166)
Illiteracy					0.564			-0.708
					(1.439)			(1.469)
Householdsize						0.461***		0.250
						(0.178)		(0.188)
Agricultural GDP Before 2005							-0.224**	-0.181*
							(0.096)	(0.093)
NREGA Dynamic Effect								
Monsoon Rain	-1.758***	-1.725***	-1.708***	-1.852***	-1.760***	-1.596***	-1.883***	-1.812***
	(0.356)	(0.348)	(0.351)	(0.333)	(0.355)	(0.373)	(0.420)	(0.428)
NREGA x Monsoon	1.457***	1.489***	1.381***	1.565***	1.459***	1.401***	1.589***	1.585***
	(0.369)	(0.372)	(0.361)	(0.360)	(0.368)	(0.373)	(0.454)	(0.464)
Observations	2886	2886	2886	2886	2886	2886	2496	2496
Number of Districts	222	222	222	222	222	222	192	192

Notes: All regressions include region-by-time fixed effects and district fixed effects. The time period is restricted to the period before NREGA was introduced. All regressions are estimated using Poisson models with the dependent variable being the number of conflict events per district and year. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10: Explaining the NREGA Effect: Monsoon Rainfall and NREGA Infrastructure Construction

	Road Construction			Land Development		
	(1)	(2)	(3)	(4)	(5)	(6)
Monsoon	0.033**	0.040*	0.059***	-0.020	-0.056**	-0.042**
	(0.014)	(0.021)	(0.017)	(0.013)	(0.025)	(0.021)
Phase 2 x Monsoon		-0.006			0.031	
		(0.030)			(0.030)	
Phase 3 x Monsoon		-0.016			0.084***	
		(0.027)			(0.030)	
LWE Affected x Monsoon			-0.056***			0.050*
			(0.022)			(0.026)
Observations	2894	2894	2894	2894	2894	2894
Number of Districts	504	504	504	504	504	504

Notes: All regressions include region-phase-time fixed effects and district fixed effects. Monsoon rain is the previous growing season's Monsoon rainfall realisation. The dependent variable is the number of violent incidences per quarter in columns (1) and (2) and an indicator whether there was any violent incidence in columns (3) and (4). Mining Sector Share is the share of the districts domestic product that is generated in the Mining sector based on data between 1998 and 2005. Standard errors are clustered at district level in column (1) and (2), while in column (3) and (4) they are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. District distances are computed from district centroids. stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A11: After the Introduction of NREGA: Reduced Form Relationship between Rainfall, Agricultural Production, Wages and Violence

	Incidence					Intensity				
	(1) Overall	(2) Phase 1	(3) Phase 2	(4) Phase 3	(5) Naxalite	(6) Overall	(7) Phase 1	(8) Phase 2	(9) Phase 3	(10) Naxalite
-8 Yrs to NREGA x log(Monsoon)	-0.035***			-0.022*	-0.126***	0.382			0.287	1.475
	(0.012)			(0.012)	(0.036)	(0.849)			(0.983)	(3.907)
-7 Yrs to NREGA x log(Monsoon)	-0.021		-0.022	-0.008	-0.064	-0.706		-1.544	0.017	1.336
	(0.018)		(0.039)	(0.021)	(0.053)	(0.981)		(1.581)	(0.848)	(1.040)
-6 Yrs to NREGA x log(Monsoon)	-0.066***	-0.091**	-0.091	-0.030	-0.186***	-1.629***	-1.336***	-3.160**	-1.489*	-2.003***
	(0.021)	(0.038)	(0.056)	(0.022)	(0.054)	(0.368)	(0.472)	(1.495)	(0.865)	(0.483)
-5 Yrs to NREGA x log(Monsoon)	-0.053***	-0.170***	-0.108	-0.004	-0.164***	-1.760***	-1.654*	-2.325***	0.178	-2.980***
	(0.018)	(0.055)	(0.082)	(0.012)	(0.058)	(0.421)	(0.853)	(0.614)	(0.548)	(0.534)
-4 Yrs to NREGA x log(Monsoon)	-0.058***	-0.160***	-0.060*	-0.020	-0.174***	-0.852*	-0.853	-1.179	0.080	-2.225***
	(0.017)	(0.058)	(0.034)	(0.015)	(0.037)	(0.472)	(0.691)	(0.812)	(0.961)	(0.533)
-3 Yrs to NREGA x log(Monsoon)	-0.046**	-0.093*	-0.078*	-0.003	-0.181***	-1.254***	-1.364***	-1.223	-0.838**	-2.237***
	(0.020)	(0.051)	(0.042)	(0.019)	(0.054)	(0.369)	(0.470)	(1.244)	(0.397)	(0.366)
-2 Yrs to NREGA x log(Monsoon)	-0.028	-0.091**	-0.081	0.028	-0.132**	-1.689***	-1.103*	-2.470***	-1.276***	-2.376***
	(0.019)	(0.044)	(0.051)	(0.020)	(0.053)	(0.336)	(0.619)	(0.798)	(0.406)	(0.394)
-1 Yrs to NREGA x log(Monsoon)	-0.066***	-0.179***	0.020	-0.052**	-0.138***	-2.071***	-2.389***	-1.656	-1.613***	-3.005***
	(0.021)	(0.053)	(0.052)	(0.024)	(0.052)	(0.303)	(0.398)	(1.384)	(0.462)	(0.430)
0 Yrs to NREGA x log(Monsoon)	-0.034	-0.080	-0.081	0.005	-0.147	-1.288***	0.029	-0.650	-2.325***	0.205
	(0.024)	(0.061)	(0.070)	(0.022)	(0.094)	(0.440)	(0.591)	(0.585)	(0.725)	(0.524)
1 Yrs to NREGA x log(Monsoon)	-0.037*	-0.057	-0.025	-0.028	-0.052	-1.011***	-0.095	-0.689*	-2.077***	0.204
	(0.022)	(0.048)	(0.050)	(0.027)	(0.060)	(0.316)	(0.483)	(0.389)	(0.607)	(0.551)
2 Yrs to NREGA x log(Monsoon)	0.008	0.030	-0.001	0.009	0.024	-0.681*	-0.575	0.566	-1.317**	0.231
	(0.021)	(0.052)	(0.042)	(0.026)	(0.070)	(0.348)	(0.556)	(0.854)	(0.553)	(0.521)
3 Yrs to NREGA x log(Monsoon)	0.022	0.117**	-0.006	-0.016	0.085	-0.128	0.542	-0.394	-0.970	1.431**
	(0.027)	(0.055)	(0.051)	(0.040)	(0.074)	(0.502)	(0.951)	(0.744)	(0.725)	(0.720)
4 Yrs to NREGA x log(Monsoon)	0.014	0.029	0.029	0.009	0.020	0.802*	1.252**	2.092*	-0.571	1.498**
	(0.028)	(0.084)	(0.084)	(0.027)	(0.066)	(0.457)	(0.610)	(1.191)	(0.773)	(0.691)
5 Yrs to NREGA x log(Monsoon)	-0.019	-0.013	-0.051		0.002	-0.195	-0.333	1.404		-0.309
	(0.053)	(0.073)	(0.078)		(0.089)	(0.376)	(0.459)	(1.197)		(0.439)
6 Yrs to NREGA x log(Monsoon)	0.070	0.043			0.094	0.916*	1.056**			1.423**
	(0.060)	(0.062)			(0.086)	(0.546)	(0.527)			(0.580)
Observations	7059	2470	1573	3016	2600	2760	1315	626	819	1621
Number of Districts	543	190	121	232	200	222	102	53	67	130

Notes: All regressions include region by NREGA phase and time fixed effects and district fixed effects and constrain the analysis to the period before NREGA was introduced. For columns (1)-(3) standard errors are adjusted to reflect spatial dependence as modelled in Conley (1999). Spatial autocorrelation is assumed to linearly decrease in distance up to a cutoff of 500 km. Poisson regressions present standard errors clustered at the district level, stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Data Appendix

B.1 Conflict Data

Empirical research on the economics of conflict almost always suffer from severe data limitations. This lies in the nature of the subject of study, that typically places that exhibit conflict are only weakly institutionalised with little official report of violence and little press and media coverage. Blattman and Miguel (2010)'s review cites that the correlation across different civil war datasets ranges from 0.42 to 0.96, which may be the reason why empirical results are often not reproducible using similar identification strategies, but different datasets or variable definitions (e.g. Ciccone (2011)).

There exists no broad conflict dataset that covers India or South East Asia as a whole. This gap was filled through the violence dataset introduced in Fetzer (2013). This paper documents the process through which in the Indian context 28,638 newspaper reports were transformed into a workable conflict dataset using both machine-learning, semi-automated coding techniques and scalable manual hand-coding methods.[54] This section sketches the semi-automated process through which the daily newspaper clippings are transformed (more details are provided in Fetzer (2013)). A typical sample may look as follows:

> Two unidentified terrorists massacred six members of a family and left a seventh injured at Mangnar Top, Poonch district, on December 31, 2001. Local residents refused to cremate the bodies of the slain victims, insisting that a Union Minister should visit the area and take notice of the increasing terrorist violence there.

The semi-automated routine defines a terrorist-incident as an Event-tuple, $E = \{L, T, V, S, O\}$ defined by a location L, a date or time of the event T, a verb V that indicates the type of violent act, and the verb's associated subject S, the perpetrator of the act and the object O that was subjected to the act V. The semi-automated routine tries to fill all these elements of the tuple for each sentence using common machine-learning algorithms implemented in natural language processing packages.

I work with the following set of Trained Natural Language Processing Algorithms:

1. Sentence Detection to break up individual sentences.
2. Semantic Role Labelling (SRL) to tag the grammatical structure of words in relation to one another.
3. Named Entity Recognition (NER) to identify names (places, institutions, names) lives off spelling, preposition and gazetteer. Complemented with dictionary of 1,978 spelling variations.

[54] The raw material was a set of 28,638 newspaper clippings collected by the Institute for Conflict Management in New Delhi through the South Asian Panel on Terrorism (SATP) since 2001, see `http://www.satp.org`, accessed in October 2012.

4. Part of Speech Tagging (POS) to tag role of words (subject, verb, object)

These are together implemented in SENNA (Collobert et al., 2011), available as open-source in C. The sample output for the above sentence would look like:

Two			B-A0	B-A0	
unidentified			I-A0	I-A0	
terrorists			E-A0	E-A0	
massacred		massacred	S-V		
six			B-A1		
members			I-A1		
of			I-A1		
a			I-A1		
family			E-A1		
and					
left		left		S-V	
a				B-A1	B-A1
seventh				I-A1	E-A1
injured		injured		I-A1	S-V
at				I-A1	B-AM-LOC
Mangnar	B-LOC			I-A1	I-AM-LOC
Top	E-LOC			I-A1	I-AM-LOC
,				I-A1	I-AM-LOC
Poonch	S-LOC			I-A1	I-AM-LOC
district				I-A1	E-AM-LOC
,				I-A1	
on				I-A1	
December				I-A1	
31				I-A1	
,				I-A1	
2001				E-A1	

In the above text-snippet, only one sentence satisfies the requirement of all elements forming an event tuple $E = \{L, T, V, S, O\}$ being present. This yields:

$$\begin{aligned} E_1 &= \{'\text{Mangar Top Poonch}','\text{December 31 2001}', \\ &\quad '\text{massacre}','\text{two unidentified terrorists}', \\ &\quad '\text{six members of a family at Mangnar Top, Poonch district}'\} \end{aligned}$$

An incident is counted as long as all pieces of information can be deduced from the underlying sentence. This is essentially mimicking the process through which humans would code this data manually. An exhaustive list of verbs is used to spot events and a sentence is normalised to contain at most one event. The individual elements of the tuple E are then transformed by assigning labels to the snippets indicating whether the actor was a terrorist, security force or a civilian and similarly for who subjected to the act V. Note that in the sentence there exists a further event:

$$
\begin{aligned}
E_2 \;=\; & \{'\text{Mangar Top Poonch}','\text{December 31 2001}', \\
& '\text{left}','\text{two unidentified terrorists}', \\
& '\text{a seventh injured at Mangnar Top, Poonch district}'\}
\end{aligned}
$$

As described in Fetzer (2013), a sentence will be counted as containing information of at most one incident. The data has been evaluated in Fetzer (2013) and correlates very well with hand-coded data. The correlation between this automatically retrieved data and the hand-coded data for the Naxalite conflict used by Vanden Eynde (2011) is at least 93%.

B.2 Comparison of Results with Global Terrorism Database

This section highlights that the results obtained in my paper can not be replicated when studying the conflict for India contained in the Global Terrorism Database (GTD) collected by National Consortium for the Study of Terrorism and Responses to Terrorism at the University of Maryland. This database has been used in more than 30 journal publications and thus, serves as an interesting testing ground. Unfortunately, the GTD database does not come at a district level spatial resolution. However, it provides the nearest big town to where the incident occurred. In order to be able to compare the datasets, I geo-code the locations of the nearest towns to obtain a similar district level count variable of the number of conflict events. I then estimate the main specifications using the number of terrorist incidences in the global terrorism database as a left-hand side. The results are presented in Table B1.

Table B1: NREGA Effect in the GTD and Fetzer (2013) dataset

	Fetzer (2013) Dataset			Global Terrorism Database		
	(1) Pre NREGA	(2) Dynamic	(3) Level	(4) Pre NREGA	(5) Dynamic	(6) Levels
Monsoon	-0.866***	-1.330***	-0.680***	-0.985	-1.338*	-1.062**
	(0.270)	(0.306)	(0.261)	(0.684)	(0.764)	(0.462)
NREGA x Monsoon		1.098***			0.359	
		(0.388)			(0.676)	
NREGA			-0.540***			-1.098
			(0.166)			(1.264)
Observations	2841	8868	10199	851	5268	5268
Number of Districts	148	217	217	57	186	186

Notes: All regressions are estimated using a pseduo-maximum likelihood estimator, whose moment conditions coincide with a Poisson model. Regressions in columns (1)-(2) and (4)-(5) include region-phase-time fixed effects as well as district fixed effects, while results for columns (3) and (6) come from a regression with time- and district fixed effects. The dependent variable is the number of incidences per district and quarter. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Columns (1)-(3) study the dataset used in this paper, while columns (4)-(6) use the GTD database. In column (4) it becomes obvious that in the GTD data, there appears to be no statistically significant correlation between rainfall and conflict, while there is a strong documented in the Fetzer (2013) data in column (1). The geographic coverage of the GTD dataset is a lot more limited before the introduction of NREGA, with only 57 districts reported as having violent incidences before NREGA was introduced while there are almost three times as many districts reported in the other datasets. The moderating effect of NREGA is seen only in column (2), but not in column (5), albeit the coefficient is positive.

As the number of districts covered in the GTD database seems to increase dramatically when expanding the analysis to the whole time-period in column (5) it becomes instructive to study how the correlation between these two datasets has evolved over time. I regress the two datasets onto one another, allowing for there to be a separate coefficient for each year:

$$GTD_{dt} = \delta_d + b_{rt} + \sum_{t=2000}^{2010} \gamma_t A_{dt} + \epsilon_{dt}$$

The estimated coefficients γ_t are plotted out in Figure B1.

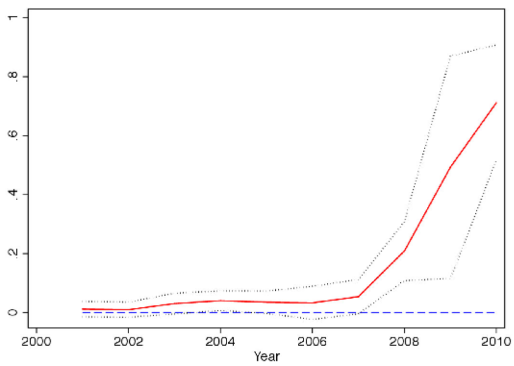

Figure B1: Relationship between Fetzer (2013) and GTD Data over Time

The specification, by using district- and region by time fixed effects takes out any fixed-conflict region and time varying reporting differences, while the district fixed effects remove any time-invariant district specific reporting biases. The coefficients paint a very stark picture: the datasets do not compare well at all before 2007. The good news is that the coefficients are consistently positive, suggesting that the overall correlation is positive. However, the point estimates are very small and only sometimes statistically significantly different from zero. This suggests that int he earlier years it is extremely unlikely for an incident captured in one dataset to appear in the other. In more recent years, the data become increasingly similar.

Why have the two datasets converged? It appears that the underlying data source in the GTD database has evolved significantly over time. Since 2008, the SATP reports feed into the GTD database, while before that the GTD database was mainly fed by newswire services. By 2010, more than 53% of the incidences in the GTD database were directly referenced with a report from the SATP newspaper clippings dataset. This is clearly, a lower bound since for many reports in the GTD dataset one can manually find references in the SATP dataset, but not necessarily vice versa.

While the level of violence reported in the GTD database seems to be significantly lower for early years, it is important for the identification whether this mismatch in reporting is correlated with rainfall realisations.

In order to explore this, I measure the differences and the absolute value of the differences between the two datasets and run the three specifications from above again.

The results are presented in table B2. The coefficients suggest that a positive rainfall realisation in the preceding month is significantly correlated with a lower reporting difference, i.e. implying that the mismatch between the Fetzer (2013) dataset and the GTD dataset is smaller. This highlights that reporting is likely to be endogenous to past weather and thus, past income realisations. While this is something that can fundamentally, not be checked, I believe that this is more likely to be a problem for the GTD database, where reporting has been found to correlate with Foreign Direct Investment in Fetzer (2013). The introduction of NREGA appears to have further reduced the mismatch between the two datasets.

Table B2: Evolution of Reporting Differences between GTD and Fetzer (2013) datasets

	Reporting Difference			Absolute Value of Reporting Difference		
	(1) Pre NREGA	(2) Dynamic	(3) Level	(4) Pre NREGA	(5) Dynamic	(6) Levels
Monsoon	-0.078**	-0.090**		-0.107***	-0.136***	
	(0.032)	(0.036)		(0.030)	(0.034)	
NREGA x Monsoon		0.051			0.060	
		(0.042)			(0.043)	
NREGA		-0.398	-0.048		-0.503*	-0.094*
		(0.269)	(0.055)		(0.278)	(0.050)
Observations	12657	25521	27693	12657	25521	27693
Number of Districts	543	543	543	543	543	543

Notes: All regressions are simple linear regressions with time- and district fixed effects. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

If we take this and the previous results together, this suggests that there is some systematic differences to the GTD dataset which correlates with rainfall in a systematic way and the introduction of NREGA may have lead to a moderation of this reporting difference. Since the two datasets appear to be converging over time and the coverage of the GTD dataset actually expanding, it seems reasonable to conclude that the SATP data source on which the Fetzer (2013) dataset is a more consistent way to measure conflict.

B.3 TRMM Rainfall Data

This paper is the first one in economics to use data from the Tropical Rainfall Measuring Mission (TRMM) satellite, which is jointly operated by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace and Exploration Agency (JAXA). The satellite carries a set of five instruments to construct gridded rainfall rates at very high spatial and temporal resolution.

The TRMM Multi-Satellite Precipitation Analysis provides daily rainfall from 1998 to 2012 at a fine spatial resolution of 0.25 by 0.25 degree grid-cell size. The data from the various instruments aboard the satellite are cleaned and calibrated using additional data from the accumulated Climate Assessment and Monitoring System (CAMS). The output of the algorithm are 3-hourly rainfall rates for that time-period. This is then scaled up to obtain monthly mean precipitation rates, which in turn are transformed into overall monthly rainfall.

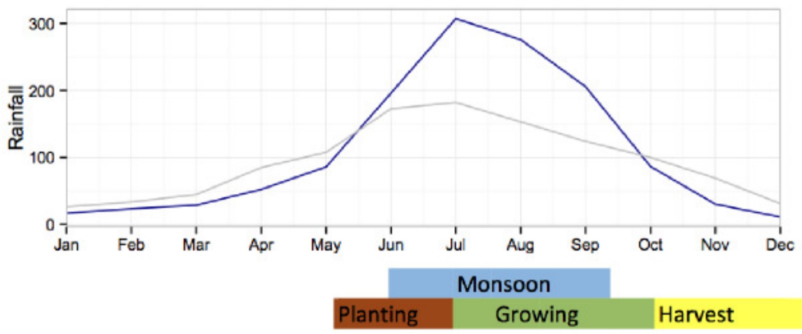

Figure B2: Rainfall and Growing Season for Andhra Pradesh

Remotely sensed weather data is an important source of data, in particular, for less developed countries, where observational data is scarce. This is particularly relevant in the case of India, where observational weather may vary in systematic ways. There are three main drawbacks. First, most observations come from rain gauges, where measurements are taken once a day. Climatologist are concerned about rain gauges in particular in tropical- or subtropical areas, since most rainfall is convective. Such convective rainfalls are highly local, generating intermittent and scattered rainfall, which may not be picked up using rain gauges, if the network is not spatially fine enough. The TRMM satellite orbits the earth every 90 minutes, thus providing multiple observations each day. An alternative is to consider data from weather radars. Rainfall radar may provide estimates for rainfall in a radius of 200 km around the station, however it is unreliable for distances in excess of 200 km. In the Indian case, rainfall radar data is not made available and would be problematic, since most reporting radar stations are clustered along the coast. The third general concern regarding observational weather data is the fact that reporting may be endogenous e.g. to violence or other variables that are correlated with the dynamics of violence. This has been highlighted recently by Smith et al. (2011), who show that Somalian piracy has generated a "black hole" in the Indian ocean, where observational weather data from merchant vessels is not available anymore, as vessels take routes avoiding piracy infested areas.[55]

I prefer the TRMM data as it is less subject to systematic measurement error, as the underlying data source is consistent over time. This is not the case with rain gauge based data, such as the GPCC as used by Miguel et al. (2004), Ferrara and Harari (2013) and Kudamatsu et al. (2014) and many others. In the case of India, the number of reporting weather stations for the GPCC data set varies from year to year. In 2001 there were a total of 1197 stations that reported at least some data, while in year 2008 that number dropped to 978. On average, 15.7 % of the district-year observations have some rainfall station reporting data. This pattern varies systematically with violence as is shown in table B3. The table presents results from the same specification as in the main part of the paper, including region-by NREGA phase time fixed effects and district fixed effects. The dependent variable is an indicator whether any station reported data for that district and year. The regressor is either an indicator whether a district experienced any violent incident in the last year (column (1)) or the number of incidents in column (2).

Table B3: Weather Station Reporting in GPCC Varies with Violence

	(1)	(2)
Any Violence	-0.013	
	(0.009)	
Attacks		-0.002**
		(0.001)
Mean of DV	.157	.157
Observations	5440	5440
Number of Districts	544	544

Notes: All regressions are simple linear regressions with time- and district fixed effects. Robust standard errors clustered at the district level are given in the parentheses with stars indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The coefficient on the violence indicator is insignificant, with a p-value of 18.5%. The coefficient on the number of attacks is significant at 5%, indicating that one additional attack per year decreases the probability of a rain gauge station reporting data in the subsequent year by 1.3% percent, when evaluating it against the mean of the

[55] Another example is the case of Vanden Eynde (2011), who had to merge several districts together in order to obtain consistent rainfall estimates, since many stations simply fail to report rainfall estimates. Most of these stations are located in places with conflict or in newly created districts or states.

dependent variable. Despite this general concern, my results are robust to using either the GPCC data (Schneider et al. (2011)) or the Indian Meterological Department data used in Vanden Eynde (2011).

B.4 Temperature Reanalysis Data

As a solution to the problem of limited data availability for ground measurements, I construct temperature readings from a gridded daily reanalysis dataset that uses remote sensing data and sophisticated climate models to construct daily temperature on a 0.75° (latitude) x 0.75° (longitude) grid (equivalent to 83km x 83km at the equator).[56] The ERA-Interim reanalysis is provided by the European Centre for Medium-Term Weather Forecasting (ECMWF).[57] As the grid is significantly coarser than the rainfall data, I construct inverse distance weighted daily mean temperatures for all grid points within 100 km of the geographic centre of each district. The weighting used is the inverse of the distance squared from the district centroid.

B.5 Agricultural Production, State Level Harvest Prices and District Domestic Product

For every district, I only consider crops that have been consistently planted on at least 1000 acres for the whole period that the state reports agricultural production to the data dissemination service of the Directorate of Economics and Statistics with the Ministry of Agriculture.[58] This leaves the following crops: bajra, barley, castor-seed, chilly, cotton, gram, groundnut, jowar, jute, linseed, maize, mesta, potato, ragi, rapeseed, rice, sesamum, sugarcane, tobacco, tumeric, tur-arhar and wheat. These capture India's most important staple crops as well as cash crops. Underrepresented is production of fruits or other horticulture products.

For each of these crops, I obtained state-level farm harvest prices to compute a district level measure of the agricultural output value. Unfortunately, district level harvest prices were not available throughout or only for a limited number of crops that did not match well with the actual planted crops. For that reason, I stuck with the state-level prices. The resulting dataset is an unbalanced panel, since not all states consistently report data to the Ministry of Agriculture information systems.

For the quantification exercise on the insurance value, I scale up the district level agricultural output value to match the district domestic product for the year 2000. The district domestic product is an estimate of local area incomes that has been produced for the period 1998-2005, but is not available for more recent years.[59] It relies on a large set

[56] To convert degrees to km, multiply 83 by the cosine of the latitude, e.g at 40 degrees latitude 0.75 x 0.75 cells are 83 x cos(40) = 63.5 km x 63.5 km.

[57] See Dee et al. 2011 for a detailed discussion of the ERA-Interim data.

[58] This data is available on `http://apy.dacnet.nic.in/cps.aspx`, accessed 14.12.2013.

[59] The data is available from `http://planningcommission.nic.in/plans/stateplan/index.php?state=ssphdbody.htm`, accessed on 21.06.2014.

of input statistics, including the Annual Survey of Industry, the National Sample Survey and Crop Production Surveys. The district domestic product construction is discussed in detail in Katyal et al. (2001). I obtain a baseline measure of the agricultural output per capita from the district domestic product. This measure will be unambiguously larger than the computed agricultural output value derived from the crop production statistics, as I only include crops that have been consistently reported for the time period that a state reports data to the Directorate of Economics and Statistics. I compute for each district a scaling factor ω_d that measures the share of the agricultural output value per capita that is captured in the agricultural district domestic product. I then simply scale up the agricultural output value per capita by this scaling factor. This preserves the variation but likely gets the agricultural output value closer to the true. This scaled agricultural output value per capita will be used for the quantification exercise to evaluate how much insurance NREGA provides.

B.6 Agricultural Wages in India

This appendix describes the process of how the agricultural wage data was cleaned and put in shape for the analysis in the paper. The data is only source providing consistently reported wage data for the whole of India. The raw data gives monthly wages for male, female and children, broken into skilled- and unskilled agricultural labour and different types of labour. The types of skilled labour are blacksmith, carpenter and cobbler, while unskilled labour combines ploughman, reaper/harvester, sower, weeder, other agricultural labour. In some states, these separate unskilled labour categories are not reported[60], but rather, a category "Field Labour Wages" is reported. This is conceived to be an average of the different categories.

In some districts these wages are reported throughout the year, while in others the wages are reported only in the parts of the year, when particular activities are actually carried out (i.e. sowing wages in the early Kharif season of May, June and July), while harvesting wages are reported in the fall of a given year.

After digitising and entering the raw data, I proceed to construct an annual level agricultural field-labour wage as my main dependent variable. For each district, there may be multiple wage-observations in case there are multiple reporting centres. I generate a balanced panel requiring each quarter of the year to have at least one non-missing observation of agricultural wages belonging to the particular category of unskilled labour. I then construct the simple average across these wage-observations.

There are advantages and disadvantages to this approach. In particular, by construction, this implies that within a year, some field labour wage observations are noisier then others. This can be taken into account by adequately weighting the observations. As an alternative, I can impose the requirement that there be at least one observation for each different unskilled labour category within a quarter. This condition is very stringent, as

[60] The states for which this is the case are Andhra Pradesh, Karnataka and Maharashtra.

it fails to recognise the types of agricultural activities that are pursued during a year. This approach reduces the number of districts significantly, but the results remain the same.

B.7 NREGA Data Sources and Roll Out

The data for the roll-out of NREGA come from the Ministry of Rural Development, which is responsible for administering the scheme. The sequence of roll-out was highly endogenous to a set of district level characteristics, such as the share of scheduled caste, scheduled tribe population, baseline agricultural productivity, literacy and existing levels of conflict. This becomes obvious when considering Figure B3. This picture highlights that a lot of districts in the east of India received NREGA in the first round. A lot of these districts did suffer from Naxalite violence. As discussed in the main body, I do not require exogeneity of treatment to levels of violence for my empirical design. There are two main sources for data on NREGA take-up. These are the district-level monthly-progress reports (MPR) and data coming from the Management Information System (MIS). The latter is a completely non-paper based system that has only become mandatory to use in the financial year but was still not fully operational until 2010-2011.

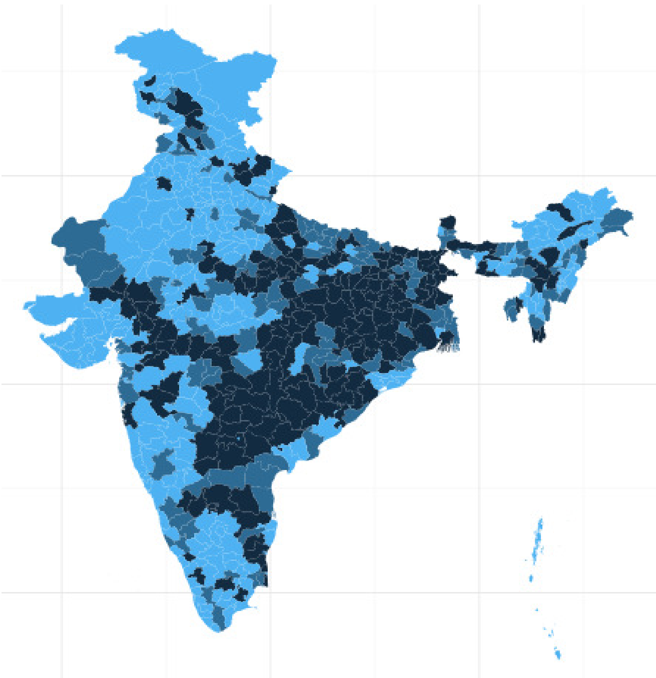

Figure B3: Phases of the NREGA Roll-out across India

There are a lot of issues regarding the reliability of either datasets, as there is quite some mismatch between the two datasets, especially in the earlier years when the MIS was introduced.[61] This may be due to partial compliance in the MIS after it had been

[61]See for example mismatch between MIS data and National Sample Survey returns data highlighted by `http://www.indiatogether.org/2013/jun/gov-nregs.htm`,accessed on 12.06.2013.

introduced, but could be also because the MPR system is more subject to manipulation. It is difficult to asses the underlying divergence in the two databases.

The MPR data is available continually from 2006 to the financial year 2010-2011, from which point onwards I rely on data from the MIS.[62] The format of the reports has changed considerably, with the major break occurring in 2011. This is partly due to the evolving nature of NREGA. Ministry of Rural Development (2009) details that several programs by the Ministry of Water Resources are to be joined with the NREGA by 2011. An important part of this program are rural sanitation projects that are funded by the Ministry of Water Resources for a set of targeted districts. This implies that there are district-specific breaks in the NREGA data. In the empirical specifications which combine data from before and after 2011, I flexibly control for these breaks by allowing the district fixed effects to be different before and after 2011.

I focus on a set of variables measuring take-up, project expenditures and overall expenditures at the district level. For the take-up I study cumulative person days provided, cumulative number of (distinct) households provided employment as well as the number of days per household at the district level. I also look at the number of person days for scheduled caste and scheduled tribe populations, as well as the share of person days that accrue to females.

For the NREGA project measures, I study the total cost or number of ongoing projects at the end of each financial year.[63] For overall expenditures, I study total expenditure in a district and year or total labour expenditures.

Despite having access to NREGA for many months in a financial year, I only study the reported metrics at the end of each financial year (that is March of each calendar year). This becomes necessary as there are significant reporting delays which induce large jumps in the cumulative month on month measures which are less likely driven by participation, but more likely due to reporting issues.

I construct the NREGA take-up, participation and project data to match the Monsoon calendar as in the main exercises.

[62] Thanks to Clement Imbert for sharing NREGA MPR data for the earliest years.

[63] The categories in the data that are consistently reported are: ”Micro Irrigation Works”,”Drought Proofing”,”Water Conservation and Water Harvesting”,”Provision of Irrigation facility to Land Owned by Scheduled Caste/ Scheduled Tribe”.