Renewable Resource Shocks and Conflict in India's Maoist Belt

Kishore Gawande (Texas A&M)

Devesh Kapur (University of Pennsylvania)

Shanker Satyanath (NYU)

ABSTRACT

Is there a causal relationship between shocks to renewable natural resources, such as agricultural and forest lands, and the intensity of conflict? In this paper we conduct a rigorous econometric analysis of a civil conflict that the Indian Prime Minister has called the single biggest internal security challenge ever faced by his country, the so called Maoist conflict. We focus on over-time within-district variation in the intensity of conflict in the states where this conflict is primarily located. Using a novel dataset of killings we find that adverse renewable resource shocks have a robust, significant association with the intensity of conflict. A one standard deviation decrease in our measure of renewable resources increases killings by 12.5% contemporaneously, 9.7% after a year, and 42.2% after two years. Our instrumental variables strategy allows us to interpret these findings in a causal manner.

Renewable Natural Resource Shocks and Conflict in India's Red Belt

(11,780 WORDS)

1. Introduction

The question of whether there is a causal relationship between shocks to renewable natural resources (such as vegetation or water) and the intensity of conflict is an important one for students of international political economy. [1] In recent years several IPE scholars have made the claim that there is such a causal link.[2] The claim is that where people are dependent on such resources for their livelihoods, adverse shocks to these resources generate incentives to fight for survival.[3] However the basis for a **causal** link between renewable resource shocks and conflict intensity has been strongly disputed, most powerfully by Gleditsch.[4] As Schwartz et al. summarize, "methodological issues underpin Gleditsch's critique… Gleditsch asserts that much environment-conflict research is methodologically unsound and fails to qualify as systematic research."[5] Among his major criticisms are a failure to address reverse causality and a failure to systematically control for alternative explanations for conflict.[6] Some methodological improvements have been made in a recent study by Theisen but fundamental concerns about endogeneity raised by Gleditsch remain unaddressed in this literature.[7]

We attempt to address the serious concerns about reverse causality and omitted variables bias that are currently obstacles to accepting the potential presence of a causal relationship by

[1] This literature differs from the literature on the resource curse, in that the latter focuses on nonrenewable resources such as oil and minerals.

[2] For example Homer-Dixon 1994; Hauge and Ellingsen 1998; Homer-Dixon 1999; Kahl 2006.
[3] Mildner et al. 2011, 157.
[4] Gleditsch 1998.
[5] Schwartz et al. 2000, 78.
[6] Schwartz et al. 2000, 78.
[7] Theisen 2011.

adopting an instrumental variables approach, and by focusing on over-time variation within districts of a single large country. In doing so we are fully aware of the trade-off between tight identification and generalizability that comes from shifting from a cross national to a single country design, but in the absence of powerful cross-national instruments for renewable resource shocks we consider this trade off to be essential to establish causality. A single country focus also serves identification because the background national context cannot be the source of omitted variables concerns.

We conduct a rigorous econometric analysis of a civil conflict that the Indian Prime Minister has referred to as "the single biggest internal security challenge ever faced by our country"[8] the so-called Naxal conflict. The Naxal conflict has largely been concentrated in a corridor of states in the East-Central part of the country, and takes the form of a Maoist insurrection with the goal of toppling the Indian government.[9] According to official Indian government data the conflict has resulted in 7862 deaths in the period 2000-2009.[10] Whereas the principal sources of politically motivated violence in India have either declined markedly (the insurgencies in Kashmir and in India's North-East) or remained unchanged at relatively low

---

[8] *The Economist* 2/25/10.

[9] The antecedents of the movement go back to a Maoist ideology inspired violent insurrection in Naxalbari, a village in the Eastern state of West Bengal in May 1967 (Chakravarti 2008). Violence spread over the next few years but it was violently put down by government forces and petered out by early 1970s. The movement's ideological currents, however, slowly spread across south-eastern, central and eastern part of the country and gradually gained strength from the early 1990s when the many splintered groups began coming together. The People's Liberation Guerrilla Army (PLGA) was founded on December 2, 2000, originally as the People's Guerrilla Army (PGA), by the then Communist Party of India–Marxist-Leninist (People's War). Following the merger of the PW and the Maoist Communist Centre of India (MCCI), on September 21, 2004, the PGA was renamed as the PLGA. For a detailed account of the origins of the Naxalite movement see Banerjee (1980); Singh (1995) provides an empathetic account of the movement until the early 1990s from the vantage point of a senior police officer.

[10] Data collated from the Annual Reports of Ministry of Home Affairs, Government of India.

levels (communal violence) over the past decade, violence related to the left wing Naxal (Maoist) movement remains the exception to this rule (Figure 1).

[Figure 1 here]

The Naxal conflict is well suited to the purpose of examining the causal links between renewable resource shocks and conflict, because it meets one of the central conditions under which the above link would plausibly hold; as we show later, large proportion of the population in the Naxal belt of states is dependent on natural resources for its livelihood. A growing literature has sought to understand the determinants of the Naxal conflict. However, aside from methodological issues, this literature suffers from a serious problem. It relies on data from the South Asian Terrorism Portal (SATP) and the National Counterterrorism Centre's Worldwide Incidents Tracking System (WITS) which measure conflict based on the study of reports in the English language press. This is quite problematic given the limited coverage of English language newspapers, the urban bias of the English language press, and the largely rural nature of the conflict. Moreover the SATP data is only from 2005. We address this gap in the literature on the Naxal conflict by creating our own conflict dataset which is additionally based on a survey of the local language press, which has far superior rural coverage. As such we are able to capture a substantially larger number of casualties than other currently available datasets.

As mentioned, our paper aims to explain within-district variation over-time in conflict intensity. This focus allows us to use demanding specifications with both district and year fixed effects, which effectively control for time invariant unobservable factors that may make districts different from each other, as well as time variant shocks that may strike all districts at any given point in time. Consistent with our goal of analyzing within-district over-time variation our core analysis focuses on the four core states of the Naxal belt, where there is substantial year to year variation in killing (at the district level) to be explained. (The core states are Bihar, Jharkhand, Chhattisgarh, and Andhra Pradesh.[11]) We instrument for a plausible proxy for over time variations in renewable resources, and can thus make causal claims relating renewable resource shocks to conflict intensity. Specifically, we use a satellite derived measure of vegetation "greenness" as a proxy for the agricultural and forest related natural resources on which a large proportion of residents of the Naxal belt rely for their livelihoods. Since vegetation is plausibly endogenous to conflict we instrument for vegetation with rainfall and find this to be a strong instrument.[12]

Our main finding is that adverse vegetation shocks have a robust, significant association with the intensity of conflict. A one standard deviation decrease in our measure of reweable natural resources increases killings by 12.5% contemporaneously, 9.7% after a year and 42.2% after two years. Our instrumental variables strategy allows us to interpret these findings in a causal manner and thus contribute to the debate on the causal relationship between renewable resource shocks and conflict.

In the next section of the paper we summarize the status of the literature on the relationship between renewable resource shocks and conflict and then describe the findings of the emerging literature on the Naxal conflict. In Section 3 we describe our causal story. The data, which are a novel aspect of our paper, and the econometric strategy, are the subject of Section 4. In Section 5 we present our main results and a range of robustness checks. Section 6 concludes.

2. Literature Review

[11] We also check for robustness when a fifth state Orissa is added.

[12] Brown (2010) uses the same NDVI vegetation index that we use when analyzing conflict in Darfur, but does not instrument for vegetation.

2.1. The debate over renewable resource shocks and conflict

Mildner et al. provide a comprehensive review of the debate over the causal relationship between natural resources and conflict.[13] Our aim here is to identify those contributions to the debate that are specifically the concerns of this paper. Mildner et al. point out that the literature can be divided into two broad categories. One focuses on the effects of renewable resource deprivation on conflict, and another focuses on the effects of the resource curse. Our paper falls in the category of the former set of concerns. Several scholars who have focused on the former concern have emphasized the threats presented by depletion of croplands, forests, water, and fish stocks to peoples' livelihoods.[14] They have argued that such adverse resource shocks can cause conflict by endangering livelihoods and forcing people to fight for their survival. Whereas authors of this school primarily rely on case studies to justify their claims, their broad claims have also found support in some statistical studies, for instance by Hauge and Ellingsen.[15]

The views espoused by this school have been criticized on several grounds. For instance, Goldstone has pointed the conceptual shortcomings in the literature, notably the failure to define variables clearly, as well as the failure to take account of alternative non-environmental explanations for conflict.[16] Especially relevant to our paper is Gleditsch's criticism from a methodological perspective.[17] In his view, case-based scholars are being inadequately systematic in their research. In particular, he has pointed out numerous failures in addressing reverse causality and in rigorously controlling for alternative explanations for conflict.

---

[13] Mildner et al. 2011.
[14] Homer-Dixon 1994 and 1999 and Kahl 2006.
[15] Hauge and Ellingsen 1998.
[16] Goldstone 2001.
[17] Gleditsch 1998.

Theisen has attempted to respond to Gleditsch's call for more systematic research with a quantitative research design.[18] He finds that Hauge and Ellingsen's claims of a major effect of natural resource shocks on conflict are likely exaggerated. Unfortunately Theisen too does not have a compelling strategy for addressing reverse causation and his study is marred by the inclusion of endogenous variables on the right hand side.

Since renewable resource shocks affect peoples' livelihoods, our paper is also related to the cross country literature on the causal relationship between income and conflict.[19] Most relevant is the paper by Miguel, Satyanath, and Sergenti which uses rainfall as an instrument for per capita GDP when studying civil conflict in Africa.[20] The big difference is that we are focusing on a level of analysis (district level) that is bereft of GDP statistics. In using a vegetation measure our paper offers a novel alternative way of addressing economic welfare at a local level in environments where citizens are reliant on agriculture/forest products for their livelihood.

2.2. The literature on the Naxal conflict

While India has had a history of radical peasant movements most observers trace the roots of the current Maoist movement to the "Naxalite" movement, which originated in 1967 as an anti-landlord peasant uprising in Naxalbari, a village in West Bengal. [21] While today in media and popular discussion, the terms "Maoist" and "Naxalite" are often used interchangeably, Maoism can be seen as the one of most radical brands of Naxalism. The 2004 merger of the People's War Group (PWG) and the Maoist Communist Centre (MCC) into the CPI (Maoist), along with the CPI (ML) Liberation's increasing focus on parliamentary politics, has rendered

[18] Theisen 2011.
[19] Collier and Hoeffler 1998, Fearon and Laitin 2003.
[20] Miguel, Satyanath, and Sergenti 2004.
[21] Singha Roy 2004.

"Maoist" the most appropriate term for today's conflict. Unlike Naxalism, which has no single doctrine, the specific ideology of Maoism can be found in the official documents of the CPI (Maoist). The Indian government officially terms these groups as "Left-Wing Extremism" (LWE).[22]

Guha argues that amongst India's principal socially marginalized groups – the Scheduled Castes (or Dalits), Scheduled Tribes (Adivasis) and Muslims – the tribal population has gained least and lost most in post-independence India, both because they are spatially isolated and their fragmentation has made them unable to articulate their grievances through the democratic and electoral process.[23] Their major problems are land alienation, denial of forest rights, and displacement by development projects and national parks and sanctuaries. The British introduced laws which independent India unquestioningly inherited, turning the stewards of the forest into subjects of the forest department. The resulting failures of the state and of the formal political system, according to Guha, have provided a space for the Maoists.

The magnitude of the problem has spawned a growing academic literature. Barooah examined which socio-economic variables explain the existence of Naxalite activity in some districts of India but not others using data from the Indian Planning Commission and South Asian Intelligence Review.[24] The dependent variable is the likelihood of violence. The main findings are that the probability of a district being Naxalite-affected rises with an increase in its poverty rate and falls with a rise in its literacy rate and that Naxalite activity in a district reduces the overall level of violent crime and crimes against women. These results are from cross sectional OLS regressions at the district level, and pool across heterogeneous regions.

---

[22] There are several careful ethnographic studies of the Maoist movement. For Bihar see Bhatia (2005) and for Jharkhand see Shah (2010). Harriss (2010) provides a good review of some of these studies.
[23] Guha 2007.
[24] Barooah 2008.

Iyer examines terrorist incidents in general, where terrorist activity includes separatist movements, communal violence, and the Naxal insurgency.[25] Using data from Global Terrorism Database 2, Rand-MIPT Terrors Database, and from Planning Commission and South Asian Intelligence Review, Iyer finds correlation between violence and poverty. Like Barooah, her findings are based on cross sectional OLS regressions at the district level.

Sen and Teitelbaum trace the history of the Maoist movement and use the World Incidents Tracking System (WITS) data on Maoist violence to examine the effects of mining.[26] They conclude that the geographical spread of Maoist movement is simply too wide to be accounted for mainly by mining activity. However, the weaknesses of their data on violence as well as mining activity, and failure to address endogeneity are potentially serious shortcomings.

Endogenity is also an issue for Hoelscher, Miklian and Vadlamannati who analyze cross-sectional data from six Indian states - Chhattisgarh, Andhra Pradesh, Orissa, Jharkhand, Bihar and West Bengal from 2004 to 2010.[27] Their data combines that from the South Asian Terrorism Portal (SATP), the National Counterterrorism Centre's Worldwide Incidents Tracking System (WITS) and the Global Terrorism Database. Using probit and negative binomial estimation techniques they find that conflict increases with forest cover, prevalence of conflict in neighboring districts, and the population share of members of scheduled castes and tribes.

Gomes combines databases from different sources on Maoist incidents/violence into one database.[28] His paper looks at landholdings and historical land institutions, and finds a strong effect of land inequality on Maoist violence. As described above this paper purports to be a

[25] Iyer 2009.
[26] Sen and Teitelbaum 2010.
[27] Hoelscher, Miklian, and Vadlamannati 2011.
[28] Gomes 2011.

district level analysis but does not include district fixed effects and also does not address endogeneity.

Van den Eynde examines the strategic choices of targets and the intensity of violence of Maoist insurgents using data from the South Asia Terrorism Portal (SATP) between 2005 and 2010.[29] Using variation in annual rainfall in a panel of district level casualty numbers he finds that negative labor income shocks increase violence against civilians to prevent them from being recruited as police informers. It also increases the number of rebel attacks against the government, but only if the rebels' tax base is sufficiently independent from local labor productivity (in this case access to key mineral resources). The casualty data used in the paper is from the SATP which, as mentioned, relies primarily on the English language press.

In a more general way our paper contributes to the statistical literature on the determinants of violence in India, for instance the work by Wilkinson on ethnic riots in India and by Jha and Wilkinson on prior military training on violence during the partition of British India.[30] Our paper adds to this literature by examining the effects of changes in renewable natural resources (namely vegetation) on violence.

3. Causal story relating renewable resource shocks to intensity of conflict

Our causal claim is that over-time, within-district variations in the intensity of conflict in the Maoist belt of states are significantly driven by shocks to renewable natural resources.[31] For this claim to be plausible we must first establish that the livelihoods of people in the region are significantly reliant on renewable natural resources. We present several facts that establish this

[29] Van den Eynde 2011.
[30] Wilkinson 2004 and Jha and Wilkinson 2012.
[31] As mentioned the core of this belt in our analysis is the four states listed at the outset of the paper where 90% of Maoist related killings occur.

basic point.

The two major renewable resources on which people in the region rely for their livelihood are agricultural land and forest land. More than two-thirds of the employment in the Maoist belt is in agriculture. In 2004 (the mid-point of the time period of our analysis), 73.6% of Andhra Pradesh's, 80.1% of Bihar's, 90.9% of Chhattisgarh's, 64.4% of Jharkhand's, and 69.2% of Orissa's workforce was employed in agriculture. This means that adverse shocks to agricultural land will plausibly affect economic welfare of numerous residents of the region.

Aside from agricultural land, a significant proportion of the population in the region relies on forests for its livelihood. The Maoist region includes a significant number of Scheduled Tribes.[32] In two of the four states in our analysis the share of tribals in the population exceeds 25%.[33] These tribals tend to live in or in the proximity of forests. In Chhattisgarh almost half of the land area in tribal districts consists of forest cover; in Jharkhand it is almost a third.[34] Tribals have close cultural and economic links with the forest and depend on forests for a significant part of their subsistence and cash livelihoods, which they earn from fuel wood, fodder, poles, and a range of non-timber forest products, such as fruits, flowers, medicinal plants and especially *tendu* leaves (used to wrap tobacco flakes and make *beeris*, a local Indian hand-rolled cigarette).

While dozens of press and academic accounts of the tribal peoples in central and eastern

---

[32] The nomenclature goes back to the 19th Century when the British expansion into the forested and hilly terrain of central India encountered culturally distinct groups living in relative geographical isolation. The first Census in 1872 categorized these communities as "Primitive Tribes". They were later termed "Backward Tribes" in the 1874 Scheduled Districts Act. The Indian Constitution in 1950 listed them in a "Schedule", granting them special protections and since then they have been known as "Scheduled Tribes".

[33] Indian Planning Commission, http://planningcommission.nic.in/data/datatable/0904/tab_119.pdf

[34] Forest Service of India. http://www.fsi.nic.in/sfr2003/forestcover.pdf. 60% of the forest cover of the country and 63% of the dense forests lie in 187 tribal districts (GOI, 2008).

India stress their dependence on forest resources for their livelihoods, a single macro-quantitative estimate of the degree of dependence is not available. We provide here the evidence that is available, which is strongly suggestive of heavy dependence. A World Bank study has found that in the state of Jharkhand fuel wood from forests supplies an average of 86 percent of energy needs while fodder from the forest provides about 55 percent of input requirements for domestic livestock for communities living proximate to forests.[35] The FAO cites a study of tribal households in Orissa which found that an average tribal family drew about one-half of its annual income from forests, 18 percent from agriculture, 13 percent from cattle and 18 percent from other employment.[36] A study of forest fringe communities in the Jhabua district of Madhya Pradesh (which is close to the Maoist belt) measured specific components of household income and subsequent dependence on natural resources (including forests), and found this dependence to be three-fourths for those in the lowest 25% income quartile and nearly two-thirds for those in the 25-50 percent income quartile.[37] Another study (in West Bengal which also neighbors the Maoist belt) found that the share of forest income in total income for tribal households ranges from 55-86%.[38]

None of the above is intended to suggest that the bulk of tribals rely on the forest to the exclusion of agriculture for their livelihoods. In his landmark work on the tribes of India, the ethnologist von Fürer-Haimendorf has categorized tribes into three categories based on their primary economic activity: food gatherers and hunters, shifting cultivators and settled farm

[35] World Bank 2006.
[36] http://www.fao.org/docrep/x2450e/x2450e0c.htm#joint%20forest%20management%20in%20india%20and%20the%20impact%20of%20state%20control%20over%20non%20wood%20f

[37] Narain, Gupta and van't Veld 2005.
[38] Das 2010.

populations.[39] What is clear from this categorization is that tribals are dependent on either the forest and/or agricultural resources for their livelihoods. This includes cases where tribals dependence on forest and agricultural resources is seasonal; a widespread pattern noted by von Fürer-Haimendorf is one where the tribals rely on forest products during much of the first half of the year, engage in subsistence agriculture during much of the second half, and supplement their income with activities like basket weaving during the monsoon season.

To sum up, we believe it is reasonable to expect adverse shocks to the renewable resources of forests and agricultural land to substantially affect the livelihoods of most residents of the Maoist belt of states. The question then is, what are our expectations about how this would result in greater intensity of conflict?

A shock to livelihoods creates incentives for people to join the Maoists as an alternative means of survival. However, we do not expect this to be immediately reflected in more intense conflict. Accounts by journalists (and there are no academic studies) stress the Maoist strategy of gradualism when inserting a new recruit into an actual combat situation.[40] Recruitment is followed by training. Recruits are taken far away from their home villages and are sometimes even sent for training to other parts of India, such as the North East, before returning to fight on familiar turf. This means that the full effects of resource shock driven recruitment on intensity of conflict should not be observed contemporaneously, but with a lag.

Another reason to expect a lagged effect is that the ability of Maoists to successfully fight the government is enhanced by support from the local population. One potential source of support is goodwill resulting from some genuine service being provided by the Maoists. The

[39] Fürer-Haimendorf 1982.
[40] Chakravarti 2008.

immediate reaction of many to an adverse economic shock is to borrow from a moneylender. Moneylenders in the Maoist belt are well known for being ruthless in demanding collateral when repayment is not forthcoming.[41] One of the major services that Maoist cadres provide is to offer protection for villagers from confiscation of assets by moneylenders.[42] This protection likely to be needed and provided, not when the loan is first taken, but rather when repayment is due. The goodwill of the local population should thus be at its highest not contemporaneously with the adverse economic shock, but after the passage of some time. This would generate incentives for Maoists to launch their most furious attacks at a lag from the adverse economic shock.

4. Econometric Strategy and Data

4.1 Econometric Strategy

The first aspect of our econometric strategy is the choice of a variable to capture renewable resource shocks in the Naxal belt. In the context of the Naxal belt such a variable should pick up differences between good and bad agricultural years and the differences between years when the forest is relatively fecund and when it is not. Both of these goals are plausibly achieved with a measure that captures the density of vegetation. Less dense vegetation should both be associated with poorer crops and with a less rich forest environment. A time varying measure of vegetation should thus pick up shocks to the main natural resources which residents of the Naxal Belt rely on for their livelihoods. Therefore, our proxy for natural resources in the Maoist region is the degree of vegetation in a district-year.

Vegetation may, of course, be endogenous to conflict. For instance the Indian government may destroy forests in places where conflict is anticipated, in order to facilitate

[41] Guha 1999
[42] Pandita 2011.

counter insurgency operations. Thus it is necessary to instrument for vegetation. We use rainfall as an instrument for vegetation, and this critical step allows us to make causal claims about the effect of natural resource shocks on conflict intensity.

In our tables we present results with and without instrumentation. The uninstrumented regressions have the following structure, with i referring to the district and t referring to the year:

$$\ln(\text{TotalDeaths})_{i,t} = \varphi \ln(\text{TotalDeaths})_{i,t-1} + \beta_1 \text{Vegetation}_{i,t} + \beta_2 \text{Vegetation}_{i,t-1} + \beta_3 \text{Vegetation}_{i,t-2}$$

$$+ \text{ Year fixed-effects } + \text{ District fixed-effects} + e_{i,t}, \qquad (1)$$

where, after accounting for the year and district fixed effects, the error term $e_{i,t}$ is identically and independently distributed normally. All our models are estimated with robust standard errors that are clustered at the district level. Since district dummies are included in the model, the coefficient estimates are based on the within-district variation in the data. The year-fixed effects account for macro-events that impact all districts in a given year, for example inflation. This is a demanding specification that takes account of observable and unobservable time invariant differences between districts (for example the share of Scheduled Tribes in a district varies very little over time and is thus absorbed by the district dummies). It also takes account of all common shocks affecting all districts in the sample at any point in time.

In our instrumental variables regressions we instrument for the vegetation variables above with three lags of rainfall in addition to contemporary rainfall. Given that there was a change towards a much more highly defined sensor to capture vegetation in 2001 our analysis begins with this year and runs to 2008 (the last year for which we have complete killings data.)

We also conduct robustness checks using the Arellano and Bond GMM technique in which lagged levels serve as instruments for lagged first differences. More details are provided in the results section.

4.2 Data

Killings

The primary data innovation in the paper is our attempt at improving on four other sources of data that have been used in the literature: The Government of India's official data from the Ministry of Home Affairs (MHA); the RAND-MITP Terrorism Incident database; data from the Worldwide Incidents Tracking System (WITS) from National CounterTerrorism Centre and data from the South Asian Terrorism Portal (SATP). The Ministry of Home Affairs compiles data on Naxal violence, dividing it into two categories – the number of casualties (deaths) and the number of violent incidents. While this data is available from 2001 onwards it is available only at the state level. The Rand-MIPT and the WITS data sets are world-wide and their Indian data is ad hoc and do a poor job at capturing Indian data reported by the non-English language press. The SATP data set has been assembled by the Institute for Conflict Management, available at their website, the South Asia Terrorism Portal. These data are based primarily on reports in the major English press, but does a considerably better job than the other two, especially in more recent years. However, this data set is only from 2005 and even after that it is not comprehensive while ours goes back to 2000.

The data set we have assembled does not draw from secondary sources. Instead for each state we examined reports from at least four distinct media sources:

i) National English press

ii) The local (or state) edition of a national English daily (if present)

iii) Local language (vernacular) dailies (at least two, except in Andhra Pradesh where we used one, Eeenadu, the largest circulation Telugu language daily in that state.

iv) Two wire services:Press Trust of India (PTI) and India Abroad News Service (IANS).

The data sources used to construct this data base of all Maoist incidents is detailed in Appendix Table 1. Further details about the dataset, including coding procedures, are provided in our online appendix (not for publication, but submitted separately for reviewers at the end of the tables).

[Appendix Table A1 here]

Each incident is geocoded at the district level. Our core measure of total casualties in incidents involving Maoists sums the number of civilian deaths, security personnel deaths and Maoist deaths for each year for each district. (We also look at these separately.) We are able to capture a substantially larger number of casualties than the widely used South Asian Terrorist portal. The extent to which other databases underestimate the actual number of deaths is evident in Figure 2. Only the SATP data is collected at the district level, but as Figure 2 indicates, it underreports considerably compared to our dataset.

[Figure 2 here]

There are two years in which our data picks up fewer killings than the aggregate data reported by the Government of India MHA numbers. Given the capacity and incentive problems of MHA data (in some cases states that reported more violence got more money to fight the Maoists), as a benchmark this data itself should be taken with caution. Is it possible that killings are over-reported by the vernacular press? This would imply that somehow the language medium would result in over-reporting. This might be if one believes that the rural press has empathy with the Maoists. But the newspapers we have used in this study – both English and vernacular – are all owned by urban capitalists, some regional capitalists and some national (none of the news sources are communist affiliated which might lead to an exaggeration of deaths). Moreover, such an argument implies that those who read/write English are less biased than those who do not, which has little basis.

We focus our attention in this paper to all districts in the four states of Andhra Pradesh, Bihar, Chhattisgarh, and Jharkhand, now known as the "red belt" for their strong association with the Maoist movement. 90% of total killings in our national data over the 2001-2008 period are from these districts.

The number of states and districts has changed over time. Chhattisgarh was carved out of Madhya Pradesh in 2000, and Jharkhand out of Bihar in 2000. To keep the unit of observation consistent across the years, we map districts into a common set of regions using the concordance constructed by Kumar and Somanathan.[43] Specifically, we use their 1991 mapping of over 600 Indian districts into 404 regions. For the four red states that are the focus of this study, we have data on 68 regions over the 2000-2008 period. (For purposes of avoiding confusion we refer to these regions as districts everywhere else in this paper.)

Rainfall

Rainfall data are from the high resolution (1° ×1° latitude/longitude) gridded daily rainfall dataset for the Indian region, kept annually since 1951 by the India Meteorological Department (IMD). The daily rainfall data are archived at the National Data Centre, IMD Pune. Rajeevan et al. of IMD Pune describe the collection of IMD data as follows: IMD operates about

[43] Kumar and Somanathan (2009, Tables 4-8).

537 observatories, which measure and report rainfall that has occurred in the past 24 h ending 0830 h Indian Standard Time (0300 UTC).[44] In addition, most of the state governments also maintain rain gauges for real-time rainfall monitoring. IMD has the rainfall records of 6329 stations with varying periods. Out of these, 537 are IMD observatory stations, 522 are under the Hydrometeorology program and 70 are Agromet stations. The remaining are rainfall-reporting stations maintained by state governments. Rajeevan et al. show that this data correlates well both spatially and inter-temporally with the VASClimo dataset, is a global gridded rainfall dataset constructed in Germany.

We match the capital cities of (districts in) the 68 regions in the four red states to nearest rain station and ascribe that rain data to the district.[45] The daily rainfall data are aggregated into annual data.

Vegetation

Satellite imagery is now widely used in the sciences to track changes in vegetation and forest cover.[46] We use the normalized difference vegetation index (NDVI) to measure annual change in vegetation for Indian districts. The NDVI data are derived from visible infrared and near-infrared data acquired from the MODIS sensor (Moderate Resolution Imaging Spectroradiometer) on NASA satellites. The NDVI index is computed as NDVI = (NIR – VIS) / (NIR + VIS), where NIR is the Near Infrared Band value and VIS is the visible light or the Red

---

[44] Rajeevan et al. 2006.
[45] Where many districts map into one region, the median district mapped to the nearest rain station.

[46] For example Myneni et al. 1998; Tucker et al. 2001; Nemani et al. 2003. In the Indian context, for example, Panigrahy et al. (2010) find dense forests denuding at 0.72% per annum in the Western Ghats, while Prabhakar et al. (2006) measure deforestation in the Himalayas.

Band value recorded by the satellite sensor.[47] The NASA site explains the computation as follows: healthy vegetation absorbs most of the visible light that hits it and reflects a large portion of the near-infrared light. Unhealthy or sparse vegetation reflects more visible light and less near-infrared light. NDVI for a given grid ranges from −1 to +1. Zero means no vegetation while values above 0.5 indicate substantial density of green leaves. NDVI is a good measure of forest cover. D'Arrigo et al show a high correlation between NDVI and direct measures of forest density from tree rings.[48]

We mapped the data from MODIS into 1° × 1° latitude/longitude grids for India. For each grid, monthly NDVI data are averaged to obtain the mean annual NDVI over the period 2001-2009.[49] They are then mapped into Indian regions, specifically into the 68 red districts as was done for the rainfall data.

Other Data

For robustness, we also report results using consumption data. The source for our consumption expenditures data is the Annual National Sample Surveys (NSS) taken over the 2000-2009 period are. Monthly per capita expenditure (MPCE) at the household level collected in these surveys are averaged using sample-proportionate-to-population weights to obtain MPCE

---

[47] http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_4.php

[48] D'Arrigo et. al. 2000.

[49] MODIS data are a substantial improvement over its predecessor, the AVHRR sensor which provided 20 years of data going back to 1980. According to Hansen et al (2003), MODIS is a significant advance due to three reasons. First is the finer instantaneous field of view of MODIS (250 and 500 m^2) as compared to AVHRR instruments (1 km^2). Second, MODIS was built with seven bands specifically designed for land cover monitoring, allowing for greater mapping accuracy due to more robust spectral signatures. It also reduces background scattering from adjacent pixels as the MODIS land bands were designed to limit the impact of atmospheric scattering. Third, 500-m red and near-infrared data, two bands important for land cover mapping, are created from averaged 250-m imagery. This resampling reduces the percent contribution of adjacency effects on 500-m pixels for these bands, allowing for improved land cover estimates. The result is a dataset that reveals far more spatial detail than previous efforts. For our purpose the great differences between the coarser-scale maps in the pre-MODIS era prevents a comparison on pre-2001 vegetation data with the MODIS data. It is a primary reason for focusing our analysis to the 2001-2008 period even though we have killings data in the pre-2000 years.

measure for each of the 68 districts annually over the period. The accuracy of our MPCE calculations was verified by matching our figures with those in the NSS summary reports (at the state level) produced by the government of India. While the consumption measure affirms our main results, we choose not to emphasize the consumption results due some technical measurement issues that have been raised by experts. The Tendulkar Commission report expands upon the difficulties of measuring poverty consistently over time from the NSS data.[50] Translating an objective poverty baseline measure based on calorific intake into a measure based on consumption spending in rupees is fraught with problems. Changes in the price of food and measurement error in changes in the quantity and quality of food consumed lead to measurement error in measuring the poverty line, based on which other measures of poverty are calculated, such as poverty headcount. Since our econometric analysis is primarily based on within-district variation in the data, measurement error in every period is less acceptable. We also want to drive the conflict literature to emphasize local knowledge and local circumstance much more than it has done in the past, as is reflected in our use of vegetation.

The literature has attempted to attribute the Maoist conflict to a number of sources including the extent of mining activity in this mineral-rich region of India, the predominance of underprivileged persons, specifically tribals (Scheduled Tribe or ST) and of people of lower caste (scheduled caste or SC) in these regions, the possibility of spillovers effects from neighboring regions, and the extent of inequality. We have made every attempt to incorporate these influences in our analysis. We measure mining of bauxite and iron ore in each district annually from Indiastat.com and compiled using the Indian Mineral Year Book annual district production figures published by the Indian Bureau of Mines, the proportion of each district's

[50] Tendulkar et al. 2009.

population that is SC and ST (from the NSS), a consumption expenditure Gini at the district level, and the number of the two closest districts that have experienced Maoist violence in the past year to capture spillover effect.

In sum, we have assembled a comprehensive data set for the 68 districts in the four red states of Andhra Pradesh, Bihar, Chattisgarh and Jharkhand annually over 2001-08. Since we incorporate dynamics and use up to three lags, our results are based on 340 observations consisting of a balanced panel of the 68 districts over the period 2004-08. Table A2 in the appendix provides panel descriptive statistics for the variables we use in our analysis.

[Appendix Table A2 here]

5. Results

In Table 1 we examine the relationship between vegetation and total deaths (our core measure of the intensity of conflict). Column 1 shows that vegetation is negatively associated with total deaths and the relationship is significant at the 1% level. The z statistic at the bottom of the column tests the hypothesis that the three coefficients sum to zero: $\beta_1 + \beta_2 + \beta_3 = 0$. If vegetation and its lags are strongly correlated, their collective significance is relevant. The sum of the coefficients is an essential input into calculating the long-run impact of a vegetation shock. Rejecting the null hypothesis, as the z statistic implies, indicates a strong negative (long-run) association of vegetation with killings. This is the first indication that vegetation is negatively associated with the intensity of conflict.

[Table 1 here]

Since our dependent variable is a count variable it is worth checking if the OLS result changes when we apply a negative binomial model.[51] Since the NB model encompasses the Poisson model (which assumes that the conditional mean of killings equals its conditional standard deviation) it allows a test of whether the Poisson model is appropriate. The parameter α tests for over-dispersion and rejects the Poisson process for killings. Column 2 shows the results with the NB model are substantively unchanged.

Column 3 offers another robustness check, this time with an Error Correction Model. Although the ECM has traditionally been used for time-series analysis of cointegrated data series, De Boef and Keele suggest the model's usefulness even when the data series are not cointegrated (but stationary, as the Maoist deaths and the NDVI series are in our case).[52] Equation (1) is easily translated into the ECM form as:

$$\Delta \ln(\text{TotalDeaths})_{i,t} = (\varphi-1)\ln(\text{TotalDeaths})_{i,t-1} + (\beta 1 + \beta 2 + \beta 3)\text{Vegetation}_{i,t-1}$$

$$+ \beta 1\ \Delta\text{Vegetation}_{i,t} - \beta 3\ \Delta\text{Vegetation}_{i,t-1} + \text{Year fixed-effects} + \text{District fixed-effects} + e_{i,t}, \quad (2)$$

where Δ is the difference operator: $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$. The ECM regresses the first difference of Maoist-related deaths on one lag of Maoist-related deaths, one lag of vegetation, the first difference of vegetation, the first difference of lagged vegetation, and district and year fixed effects. The immediate impact of a 1-unit vegetation shock is a β1 per cent change in Maoist-related deaths. The long-run impact of a 1-unit vegetation shock is a (β1 + β2 + β3) per cent change in Maoist-related deaths. The third column in Table 1 reports estimates from the ECM.

[51] The link function in the NB model is log-linear, so the OLS model in logs produces coefficients comparable to those from the NB model.

[52] De Boef and Keele 2008.

The immediate short-run impact of a negative vegetation shock of 0.01 is estimated to increase Total Deaths by $-13.99 \times 0.01 = 14\%$ in the distributed lag model (1) and by $-14.90 \times 0.01 = 14.9\%$ in the ECM. The long run impact of the same shock is estimated to be $(0.01) \times (-13.99 - 7.86 - 7.83) / (1-0.08)$ or a 32.3% increase in Total Deaths using (1), and $(0.01) \times (-29.67 / (.920)$ equal to (the same) 32.3% increase in Total Deaths using the ECM. [53]

Time series scholars are concerned that panels with short T (less than 15) may be afflicted with Nickell bias when fixed effects and the lagged dependent variable are simultaneously included in the same specification.[54] In columns 4 and 5 we thus check if the results change when the lagged dependent variable is dropped from the column 1 and 2 specifications and find that there is no substantive difference. The Arellano Bond GMM approach offers an alternative technique to address Nickell bias. Our results are robust to the use of the Arellano-Bond technique (not shown).[55]

Vegetation may be endogenous to conflict, for example if the police clear forests for ease of fighting. If shocks to killings are negatively correlated with vegetation because forests are denuded to facilitate counter-insurgency, then the negative coefficients on vegetation in Table 1 may be the result of this endogeneity bias. In order to address this concern we resort to an instrumental variables strategy, using rainfall as an instrument for vegetation. Table 2 shows the strong relationship between rainfall and vegetation, affirming the case for rainfall as a good instrument for vegetation. The Kleibergen-Paap Wald statistic of 10.19 indicates no weak instruments problem. Specifically, the K-P statistic implies the (small-sample) bias in the 2-

[53] The Akaike Information Criterion justifies the use of a model with two lags (results available on request).
[54] Nickell 1981.
[55] The Arellano-Bond results are available from the authors. They mirror the IV results we present. Even though the Arellano-Bond models are in first differences, and hence capture changes in the variables, their similar results are a strong robustness check of our core OLS-IV and NB-IV results.

stage estimates is less than 5% of the bias in the corresponding OLS estimates in Table 1.[56] As the table shows, there is generally a positive association between lagged rainfall and vegetation. The only significant negative coefficients are for forward variables --contemporaneous and once lagged rain in the twice lagged rain regression -- which is explained by mean reversion because of which we see good rain years following bad rain years.

[Table 2 here]

Table 3 shows our core instrumental variables second stage results.[57] In column 1 we report the results for a linear instrumental variables regression (OLS-IV). The strong negative association between vegetation and total deaths remains even with instrumentation. The point estimates are much larger, which is consistent with our instruments addressing measurement error with respect to vegetation. Essentially, the IV results exploit only exogenous variation in vegetation, which allow us to make causal statements. In column 2 we undertake the same exercise, but with a negative binomial model to better account for the count nature of our dependent variable. The results remain substantively unchanged, and are stronger in the same direction as the OLS-IV results.

[Table 3 here]

The strong effect of twice lagged vegetation on killing is consistent with the dynamic causal story described in Section 3. An adverse vegetation shock gives Naxal leaders the opportunity to recruit fighters into training camps. The recruitment is not overnight. Maoist cadres come in and undertake a variety of actions from providing rudimentary medical services to income augmenting schemes like offering protection against money lenders and forest

---

[56] Stock and Yogo 2004.
[57] All estimation is done using Stata 11. The 2SLS results use code due to Baum et al. 2007.

officials.[58] This means that the full effects of recruitment and training are not observed contemporaneously with the adverse vegetation shock, but rather after a lag. The long-run result, however, is a large increase in Maoist-related deaths.

The quantitative implications of the IV estimates bear this out. The OLS-IV results indicate that an adverse shock that reduces vegetation by 0.011 (a one within-standard deviation change – see table A2) increases killings by 12.5% contemporaneously, 9.7% after a year and 42.2% after two years. While the contemporaneous and the first lag effects are measured imprecisely, the second lag effect is statistically significant at 1%. The sum of the three coefficients is statistically significant at 1% (indicated by the z statistic towards the bottom). The long-run impact of an adverse shock that reduces NDVI by a one (within) standard deviation in a single period, is an increase of 64% in total deaths. While not reported, an error correction model (ECM) confirms the results.[59]

The estimates from the NB model in column 2 are notable as well. A one within-standard deviation decrease in NDVI increases killings by 16.5% contemporaneously, 27.7% after a one-year lag and 76.1% after a two-year lag. Columns 3 and 4 of Table 3 show that the results are substantively unchanged if we drop the lagged dependent variable from the column 1 and 2 specifications (to address Nickell bias concerns).

Instrument diagnostics are reported at the bottom of table 3. The Hansen J-statistic cannot reject the joint null hypothesis that (i) the instruments are uncorrelated with the error

[58] Pandita 2011.

[59] We estimate the ECM it in two stages, first predicting the three vegetation variables, and then estimating the ECM in (2) after replacing the lagged and differenced vegetation variables with their instrumented counterparts. The estimates from the ECM replicate the immediate and long-term impact of a vegetation shock in Table 3. The second lag of vegetation is statistically significant, confirming its importance. The long-run impact of a negative vegetation shock of 0.011 is estimated to increase Total Deaths by $(0.011) \times (-58.37 / (.923) = 63.2\%$. The ECM results are available from the authors.

term, and (ii) the instruments (rainfall) are correctly excluded from the estimated equation. The Anderson-Rubin (A-R) statistic tests the significance of the endogenous regressors. The statistic rejects the joint null hypothesis that the coefficients of the endogenous regressors in the structural equation (the three vegetation measures) are jointly equal to zero.[60] This statistic is important in assessing the results when instruments are weak, as we will see below. The rejection of the null is not surprising here since we expect rainfall to be a good instrument for vegetation.

What is the effect of successive years of vegetation depletion on violence in Maoist areas? This is a relevant question to ask because press reports indicate that construction, mining, and population pressures have all contributed to the degradation of forest and agricultural land in India. We investigate the impact of vegetation depletion by interacting the three lagged vegetation variables Vegetation_t, Vegetation_{t-1}, and Vegetation_{t-2}. The interactions are highly correlated with the individual lagged measures, and including a full set of variables produces collinear models whose coefficient estimates are not individually informative.[61] In the last three columns of Table 3 we present results from the set of interactions that are informative. The vegetation interactions are instrumented with the three rainfall lags and their interactions.

---

[60] The AR statistic actually tests the *joint* null hypothesis that (i) the coefficients of the endogenous regressors in the structural equation are each equal to zero, and that (ii) the overidentifying restrictions are valid. Thus, the AR statistic can reject either because the endogenous regressors are significant (here $\beta_1=0$, $\beta_2=0$, and $\beta_3=0$) or because the instrument orthogonality (with the error) conditions fail. A variety of tests have been proposed (Kleibergen 2002; Moreira 2003) that maintain the hypothesis that the instruments are valid. Moreira shows that the AR test is optimal when the equation is just identified (number of excluded instruments equals the number of endogenous variables), which is the case here. A caveat is that with extremely weak instruments and strong endogeneity of the regressors, the power of the A-R test (its ability to not falsely reject) may be poor.

[61] The partial correlation (after partialing out district and year fixed effects) of Vegetation_t and the interaction (Vegetation_{t-1}, $\times\text{Vegetation}_{t-1}$), for example, is over 0.75. Including all possible interactions and the linear terms terms leads to large coefficients and standard errors, with opposite signs.

If the interactions have negative coefficients, it implies that successive years of depletion only worsen violence. The first column shows that Vegetation$_{t\text{-}2}$ and its interaction with Vegetation$_t$ and Vegetation$_{t\text{-}1}$have negative coefficients, although none of them are statistically significant. This is likely due to the collinearity among the three variables.[62] Dropping the linear term Vegetation$_{t\text{-}2}$, in the next columns yields negative signs on both interaction terms and a statistically significant estimate on Vegetation$_{t\text{-}2}$ × Vegetation$_{t\text{-}1}$. This implies that the marginal effect of an increase in 2-period lagged vegetation, or ∂ Total Deaths / ∂ Vegetation$_{t\text{-}2}$, equals (−36.89 × Vegetation$_t$) – (84.83 × Vegetation$_{t\text{-}1}$). Evaluated at their means, the marginal effect is −49.94, implying that a one-sd increase (of 0.011) in 2-period lagged vegetation decreases total killings by 55%.

The last column shows results from one-lag interactions of Vegetation$_{t\text{-}2}$ with Vegetation$_{t\text{-}1}$ and Vegetation$_{t\text{-}1}$ with Vegetation$_t$. The negative and statistically significant coefficient on Vegetation$_{t\text{-}2}$ × Vegetation$_{t\text{-}1}$ indicates that the impact on killings due to the denuding of vegetation two periods ago (t-2) is strongly exacerbated if the vegetation is not renewed in the following year (t-1). The marginal effect ∂ Total Deaths / ∂ Vegetation$_{t\text{-}2}$, equals −45.92 and is precisely measured. It affirms the finding above.

The causal story we favor is the standard one in the resource shock-conflict literature, in which adverse resource shocks reduce the opportunity cost of fighting for rebels. Given this causal path it is important to check that the results are not driven by killings by security forces rather than rebels. To check this we disaggregate total deaths into civilian casualties, Maoist (militants) casualties, and security (police) casualties. Table 4 presents instrumental variables results for these categories using the same IV strategy as for the first two columns of Table 3.

[62] They are, however, informative about the marginal effect (see below).

The negative relationship between vegetation and casualties is in evidence for all three types of casualties. We also do not see a robust difference in the responses to vegetation of killings of rebels vs. security troops. (The slight difference in timing in the OLS-IV does not carry over into the negative binomial IV regressions.) Our first-hand experience with gathering killings data indicates civilian victims are sometimes classified as Maoists (especially when the incident is reported by the security forces). Combining the two categories yields very similar results.[63]

[Table 4 here]

One potential concern is whether the vegetation results reflect aspects of the technology of fighting? Less vegetation could change the balance of power between security forces and rebels by making it harder for rebels to conduct insurgency operations from the forest because it is harder for them to remain hidden as they attack or retreat, and therefore in conducting guerilla operations. In addition, thinner vegetation should make it easier to conduct for security forces to conduct counterinsurgency operations and kill rebels because the latter will find it harder to hide. If either of these mechanisms is driving the results, vegetation should be *positively* associated with security deaths minus rebel deaths. The OLS-IV results in the column labeled "ln(Sec) – ln(Mao)" in Table 4 show that this is not the case and this increases our confidence that our vegetation results are likely not picking aspects of the technology of fighting.[64]

We have mentioned the measurement issues that affect household consumption expenditure data collected annually by the national sample surveys (NSS). Regardless, we think it is useful to check whether our core results are robust to using monthly per capita consumption expenditures (MPCE) as measures of livelihood and poverty in place of vegetation. The first

[63] These results are available from the authors.
[64] There is no reason to run a negative binomial model here because the left hand side can take values of less than zero.

four columns of Table 5 present coefficient estimates corresponding to the non-IV estimates in Table 1. The results are substantively very similar and support inferences from the vegetation results. The ECM indicates that the long-run impact of an adverse shock of 136 rupees (approximately one within-sd) to monthly income increase Maoist-related deaths by $(-1.515 \times 0.136)/(1-.0897) = 23.3\%$. We take this as an affirmation of our main findings from the vegetation data that poverty is inversely correlated with Maoist-related killings.

[Table 5 here]

The IV results are reported in the last two columns of Table 5. The K-P statistics indicate that the consumption measure instruments weakly with rainfall (which is consistent with the measurement error critique in the Tendulkar Commission report mentioned earlier). However, the Anderson-Rubin *p*-value gives us a way of producing estimates that take account of the weak-instrument problem.[65] While the individual coefficients on consumption are insignificant the Anderson-Rubin *p*-value of .008 confirms that the variables are jointly robust even after standard errors are adjusted to take account of the extent of weakness of the instruments.

We now consider the influence of several key control variables on our core regressions. These variables have appeared in other studies and in popular stories and reports of the conflict. Mining contracts sold by the state to firms for bauxite and iron ore mining that rampantly denude forest lands and displaces its denizens, have been blamed as a source of conflict. Tribals and the lower caste population have been especially affected by such actions, and districts with greater

[65] The small Kleibergen-Paap statistics in Table 3 indicate that rainfall is a weak instrument for consumption. The problem of weak instruments has deservedly attracted much attention, for strong instruments are hard to find. "Weak-instrument robust" tests, beginning with Anderson and Rubin (1949, A-R), seek to produce valid inference about endogenous regressors and confidence intervals for their coefficients despite the weakness of instruments (see also Kleibergen 2002 and Moriera 2003). Technically, the Anderson-Rubin statistic "AR" defines is a quadratic form that is a function of the IV estimates **b**: q(**b**) = AR. Inverting this function produces interval for **b**. If the intervals preserve the signs for **b** then they are weak-instrument robust estimates. The A-R statistic of 3.55 rejects the null hypothesis that the coefficients on the three instrumented consumption variables are jointly zero, at the 1% level.

proportion of SC and ST populations are likely to experience greater displacement and therefore propensity to violence as they are recruited into the Maoist fold. Income inequality may be a uniting force among insurgents and districts with greater inequality have been shown to experience greater intensity of conflict in Nepal's Maoist insurgency.[66] Finally, the incidence of conflict in neighboring districts may spill over into conflict in the present district.

[Table 6 here]

Table 6 examines the robustness of our primary results to the inclusion of these influences as control variables. The first five columns include each of these controls, one variable at a time in the OLS-IV models. The spillover effect is statistically significant and indicates that if one of the closest two neighboring districts experiences Maoist violence, then killings increase 26.4%. While there is evidence that SC, inequality and mining all increase killings, that evidence is weak since the coefficients are statistically not significant. The sixth column pools these controls into one specification. Two results are notable. The first is that core results are unchanged with the addition of these controls. Second, the proportion of the ST population is positively associated with more deaths. An increase of .01 in the tribal population (on account of migration, for example), results in a 1.606% increase in killings.

[Table 7 here]

Could our results be driven by states in which Scheduled Tribes have not achieved representation in the set of parties that are competing in state politics? The states where there is such representation are Jharkhand and Chhattisgarh, which were carved out as separate states in November 2000 precisely to ensure greater representation of tribals. Ironically the Maoist insurgency surged in these areas after the new states were created and the tribals got more

[66] Nepal, Bohara and Gawande 2011.

political representation. In Table 7 we interact the vegetation variables and state dummies, with Andhra Pradesh as the baseline (excluded) state. The *z* statistics at the bottom summarize the results by summing the coefficients for each state across the three coefficients for each state (contemporaneous plus two lags). These coefficients test the hypothesis that the state's "total effect", that is, the sum of its coefficients relative to the Andhra Pradesh sum, equals zero. The *z*-statistics for the OLS_IV regression indicates that the total effect of vegetation is *negative* in all states. The only state for which the total effect of vegetation is insignificant in the OLS-IV regression is Chhattisgarh. (The *z*-statistic in the Negative Binomial regression is positive, albeit insignificant.) We verified that the effect for Chhattisgarh is significantly lower than for Jharkhand where representation of Scheduled Tribe is highest. This suggests that lack of political representation is likely not the reason for the differences across states.

There is a plausible explanation for why conflict intensity in Chattisgarh is less sensitive to vegetation shocks than in other states in the Naxal Belt. This is the one state where potential Maoist recruits have been deliberately moved by the government, away from forest and agricultural lands and into guarded camps along the highway. Since the effect of this measure is that the livelihoods of numerous potential recruits in Chhattisgarh are no longer subject to variations in vegetation, it makes sense that one does not observe a strong vegetation-killings link in this state.

While not affecting our claims with respect to the hard core Naxal Belt states it interesting to assess the effects of slightly expanding the definition of the Naxal Belt. The most logical expansion is to Orissa because, while not traditionally considered to be part of the Naxal Belt, parts of Orissa were subject to Naxal conflict. We coded killings in Orissa and checked if our core Table 3 results were affected by the inclusion of Orissa. We found that our core results were unchanged (not shown).

Finally, could the results be driven by violations of the exclusion restriction? One possible violation highlighted by Sarsons in the context of riots all over India lies in the need to account for the building of large dams in our analysis.[67] This may be a relevant consideration for an analysis which includes western India (where over two thirds of India's large dams are present and where most of India's dam building occurs). However, it is not a serious consideration for our analysis because there is no over-time variation for our region of analysis in the building of large dams. (Using the standard definition of 50 meters, no new large dams were built in our region after 2000.)

Another exclusion-restriction-related consideration is that fighters may not want to fight when it is raining, so rainfall may be directly affecting conflict via this mechanism. This violation is hard to measure. Whether or not this biases against our results and whether the bias is large enough to kill our result thus remains an unknown. However, we do note that this is not likely to be a serious factor affecting our analysis because our dataset indicates that the most killings occur outside of the rainfall season in good and bad rainfall years alike.

6. Conclusion

We demonstrate that even after addressing endogeneity concerns and subjecting our findings to multiple robustness checks there is a strong, and substantively large relationship between adverse resource and the intensity of conflict in India's Naxal belt. Aside from contributing to a central debate in the civil conflict literature, our results carry implications for policy. One approach would be to help citizens deal with adverse income shocks that result from

[67] Sarsons 2011.

adverse resource shocks. The creation of the National Rural Employment Guarantee Act (NREGA) could offer promise as an insurance against income shocks. (NREGA only came into effect midway through our period of analysis, so we cannot offer any firm statistical results along this line.) Another possibility is implementing a catastrophe insurance scheme, when the income shocks are spatially correlated, for instance due to extreme weather conditions. However, to the extent that these options rely on the same government machinery that has, through acts of commission, exacerbated the plight of tribals, their prospects may be questionable.

Our paper also suggests an alternative path that relies less heavily on the government's technical capability. Giving tribals greater access to forests and a range of forest products, whose consumption is the only available option during times of distress, can provide them with a critical self-insurance mechanism. Tribals have been denied this access for many years, especially after passage of the Forest Conservation Act of 1980. Reversing this situation was the goal of the Scheduled Tribes and Other Traditional Forest Dwellers (Recognition of Forest Rights) Act 2006, which came into force on 1st January 2008. Our paper suggests that rigorous implementation of the government withdrawal called upon by this act may offer great promise.

BIBLIOGRAPHY

R. D. D'Arrigo, C. M. Malmstrom, G. C. Jacoby, S. O. Los and D. E. Bunker. 2000.Correlation between maximum latewood density of annual tree rings. and NDVI based estimates of forest productivity. *International Jopurnal of Remote Sensing* 21: 2329–2336.

Brown, Ian A. 2010. Assessing Eco-Scarcity as a Cause of the Outbreak of Conflict in Darfur: A Remote Sensing Approach. *International Journal of Remote Sensing* 31 (10), 2513-20.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2007. "*ivreg2*: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML, and k-class regression." Boston College Department of Economics, Statistical Software Components S425401.Downloadable from http://ideas.repec.org/c/boc/bocode/s425401.html.

Banerjee, Sumanta. 1980. *In the wake of Naxalbari: A history of the Naxalite movement in India* (Calcutta: Subarnarekha, 1980).

Bhatia, Bela. 2005. The Naxalite Movement in Central Bihar. *Economic and Political Weekly*, Vol.XL: 1536-43.

Bhattacharya, Prodyut, Lolita Pradhan and Ganesh Yadav. 2010. Joint forest management in India: Experiences of two decades. *Resources, Conservation and Recycling*. 54: 469-480.

Nepal, Mani, Alok K. Bohara and Kishore Gawande. 2011. More Inequality, More Killings: The Maoist Insurgency in Nepal. *American Journal of Political Science* 55: 886–906.

Bohlken, Anjali Thomas and Ernest Sergenti. 2010. Economic Growth and Ethnic Violence: An Empirical Investigation of Hindu Muslim Riots in India. *Journal of Peace Research* 47 (5), 589-600.

Borooah, Vani K. 2008. Deprivation, Violence, and Conflict: An Analysis of Naxalite Activity in the Districts of India *International Journal of Conflict and Violence* 2(2): 317-333.

Chakravarti, Sudeep. 2008. *Red Sun: Travels in Naxalite Country.* New Delhi: Penguin Viking Press.

Collier, Paul and AnkeHoeffler.1998. "On Economic Causes of Civil War," *Oxford Economic Papers* 50.

Collier, Paul and Anke Hoeffler. 2001. "Greed and Grievance in Civil War," *World Bank Policy Research Paper* 2355 (May).

Collier, Paul and AnkeHoeffler.2002. "On the Incidence of Civil War in Africa", *Journal of Conflict Resolution* 46(1).

Das, Nimai. 2010. Incidence of forest income on reduction of inequality: Evidence from forest dependent households in milieu of joint forest management," *Ecological Economics*, Volume 69 (8): 1617–1625.

De Boef, Suzanna and Luke Keele. 2008. “Taking Time Seriously”. *American Journal of Political Science* 52(1): 184-200.

Dube, Oeindrila, and Juan Vargas. (2010). “Commodity Price Shocks and Civil Conflict: Evidence from Colombia”, unpublished working paper, NYU.

Eynde, Oliver Vanden. 2011. Targets of Violence: Evidence from India's Naxalite Conflict. http://personal.lse.ac.uk/vandeney/Targets_of_Violence.pdf

Fearon, James and David Laitin. 2003. “Ethnicity, Insurgency and Civil War.” *American Political Science Review*, 97(1): 75-90.

Gadgil and Guha, 1994. Ecological conflicts and the environmental movement in India *Development and Change*, 25: 101-136.

Gallo, Kevin P.; Ji, Lei; Reed, Brad; Dwyer, John; and Eidenshink, Jeffrey. 2004. "Comparison of MODIS and AVHRR 16-day normalized difference vegetation index composite data" *Geophysical Research Letters* 31: L07502.

Gleditsch, Nils-Petter. 2008. Armed Conflict and the Environment: A Critique of the Literature. Journal of Peace Research 35 (3), 360-380.

Goldstone, Jack A.. 2001. Democracy, Environment, and Security in Environmental Conflict, ed. Paul Diehl and Nils-Petter Gleditsch, 84-108. Boulder: Westview.

Gomes, Joseph Flavian. 2011. The Political Economy of the Maoist Conflict in India: An Empirical Analysis. Available at: http://www.uclouvain.be/cps/ucl/doc/core/documents/Gomes.pdf

Government of India. 2008. "Development Challenges in Extremist Affected Areas: Report of an Expert Group to Planning Commission." Report, April, Planning Commission, New Delhi. http://planningcommission.gov.in/reports/publications/rep_dce.pdf

Guha, Ramachandra. 2007. Adivasis, Naxalites and Indian Democracy. *Economic and Political Weekly*, August 11: 3305-3312.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg. 2003. Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm *Earth Interactions* 7, Paper No. 10: 1-15.

Hauge, Wenche and Tanja Ellingsen. 1998. Beyond Environmental Security: Causal Pathways to Conflict. Journal of Peace Research 35 (3), 299-317.

Homer-Dixon, Thomas. 1994. Environmental Scarcities and Violent Conflict: Evidence from Cases. International Security 19 (1), 5-40.

Homer-Dixon, Thomas. 1999. Environment, Security, and Violence. Princeton: Princeton University Press.

Harriss, John. 2010. The Naxalite/Maoist Movement in India: A Review of Recent Literature ISAS Working Paper, No. 109 – 08 July.

Hidalgo, F. Daniel, Suresh Naidu, Simeon Nichter, and Neal Richardson.2010. Economic Determinants of Land Invasions *Review of Economics and Statistics*, 92(3), 505-523.

Hoelscher, Kristian, Jason Miklian and Krishna Chaitanya Vadlamannati. 2011. Hearts and Mines: A District-Level Analysis of the Maoist Conflict in India. Available at: http://www.uni-heidelberg.de/md/awi/professuren/intwipol/india.pdf

Jenkins, Stephen P. and Philippe Van Kerm. 2009. The Measurement of Economic Inequality in Brian Nolan, Wiermer Salverda and Tim Smeeding (eds.) *Oxford Handbook on Economic Inequality*. New York: Oxford University Press: 40-71.

Jha, Saumitra and Steven Wilkinson. 2012. Veterans, Organizational Skill and Ethnic Cleansing: Evidence from the Partition of South Asia
Stanford Graduate School of Business Working Paper No 2092.

Kahl, Colin. 2006. States, Security, and Civil Strife in the Developing World. Princeton: Princeton University Press.

Kleibergen, F. 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70: 1781–1803.

Kumar, Hemanshu andRohini Somanathan. 2009. Mapping Indian Districts Across Census Years 1971-2001 *Centre for Development Economics, Delhi School of Economics*, Working Paper 176.

Leamer, Edward E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley.

Mehta, J. and Venkatraman, S. 2000. Poverty Statistics: Bermicide's Feast, *Economic and Political Weekly*, 35: 2377-2381.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. Economic Shocks and Civil Conflict: An Instrumental Variables Approach *Journal of Political Economy*, 112(4), 725-753.

Miguel, Edward and Shanker Satyanath. 2011. Re-examining Economics Shocks and Civil Conflict. *American Economic Journal, Applied Economics* 3(4), October 2011.

Mildner, Stromy-Annila, Gitta Lauster, and Wiebke Wodny. 2011. Scarcity and Abundance Revisited: A Literature Review on Natural Resources and Conflict. International Journal of Conflict and Violence 5 (1), 155-172.

Ministry of Home Affairs. 2006. *Annual Report 2005-2006* (New Delhi: Govt. of India).

Moreira, M. 2003. A conditional likelihood ratio test for structural models. *Econometrica* 71: 1027–1048.

Myneni, R., C. Tucker, G. Asrar, and C. Keeling.1998.Interannual variations in satellite-sensed vegetation index data from 1981 to 1991, *Journal of Geophysical Research* 103(D6), 6145–6160.

Narain, Urvashi, Shreekant Gupta, and Klaas van ’t Veld. 2005. Poverty and the Environment: Exploring the Relationship between Household Incomes, Private Assets, and Natural Assets Resources for the Future Discussion Paper 05-18.

Nemani, R. R., C. D. Keeling, H. Hashimoto, W. M. Jolly, S. C. Piper, C. J. Tucker, R. B. Myneni, and S. W. Running. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999, *Science* 300: 1560– 1563.

Nickell, Stephen John. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49 (6): 1417-26.

Nillesen, Eleanora and Philip Verwimp.2010. Grievance, Commodity Prices and Rainfall: A Village Level Analysis of Rebel Recruitment in Burundi *HiCN Working Paper* 58. http://www.hicn.org/papers/wp58.pdf

Panigrahy, Rabindra K.,; Kale, Manish P., Dutta, Upasana, Mishra, Asima, Banerjee, Bishwarup Singh, Sarnam. 2010. Forest cover change detection of Western Ghats of Maharashtra using satellite remote sensing based visual interpretation *Current Science* 98 (05)

M. Poffenberger, B. McGean (Eds.). 1996. *Village Voices, Forest Choices: Joint Forest Management in India*, Oxford University Press, New Delhi.

Prabhakar, R, Somanathan, E., Mehta, Bhupendra Singh. 2006. How degraded are Himalayan forests? *Current Science* 91 (01).

M. Rajeevan, Jyoti Bhate, J. D. Kale and B. Lal. 2006. "High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells." *Current Science* 91, 2006: 296-306.

Sarsons, Heather. 2011. Rainfall and Conflict. Manuscript.http://www.econ.yale.edu/conference/neudc11/papers/paper_199.pdf

Schwartz, Daniel M., Tom Delgiannis, and Thomas Homer-Dixon. 2000. The Environment and Violent Conflict: A Response to Gleditsch's Critique and Some Suggestions for Future Research. *Environmental Change and Security Project Report*, 6 (Summer), 77-94.

Sen, Rumela and Emmanuel Teitelbaum. 2010. Mass Mobilization and the Success of India's Maoists http://web.gc.cuny.edu/dept/rbins/conferences/RBFpdf/Sen-TeitelbaumMaoists.pdf

Shah, Alpa. 2010. *In the Shadows of the State: Indigenous Politics, Environmentalism and Insurgency in Jharkhand, India*. Durham and London: Duke University Press.

Shah, Amita and Sajitha O.G. 2009.Dwindling forest resources and economic vulnerability among tribal communities in a dry/sub-humid region in India. *Journal of International Development* 21: 419–432

Singh, Prakash. 1995. *The Naxalite Movement in India* New Delhi: Rupa and Co.

SinghaRoy, Debal K.. 2004. *Peasant Movements in Post-Colonial India: Dynamics of Mobilisation and Identity* New Delhi: Sage.

Stock, James H. and Motohiro Yogo. 2004. "Testing for Weak Instruments in Linear IV Regression." In D. W. K. Andrews and J. H. Stock(eds.), *Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*. Cambridge: Cambridge University Press.

Teitelbaum, Emmanuel. "Political Representation and Rural Insurgency in India". Working Paper.

Tendulkar, Suresh D., R. Radhakrishna, and Suranjan Sengupta. 2009. "Report of the Expert Group to review the Methodology for Estimation of Poverty". Government of India Planning Commission.

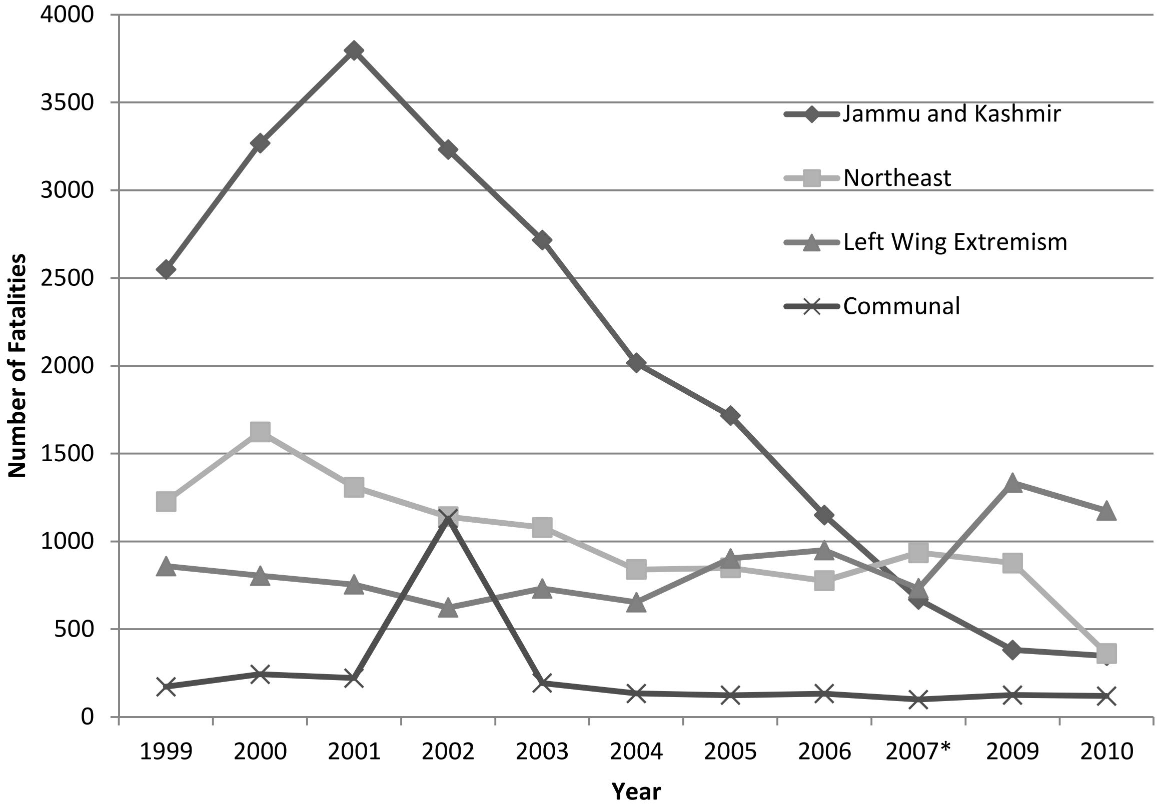
Theisen, Ole Magnus. 2008. Blood and Soil: Resource Scarcity and Armed Conflict Revisited. Journal of Peace Research 45, 801-818.

Tucker, C. J., D. A. Slayback, J. E. Pinzon, S. O. Los, R. B. Myneni, and M. G. Taylor. 2001. Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999, *International Journal of Biometeorology* 45: 184– 190.

Wilkinson, Steven 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India.* Cambridge: Cambridge University Press.

World Bank. 2006. “Unlocking the Opportunities for Forest-Dependent People in India. Washington D.C.: World Bank. Report No. 34481-IN.

Figure 1. Sources of Violence in India 1999-2010



Source: (i) Ministry of Home Affairs Annual Reports (various years), (ii) Indiastat

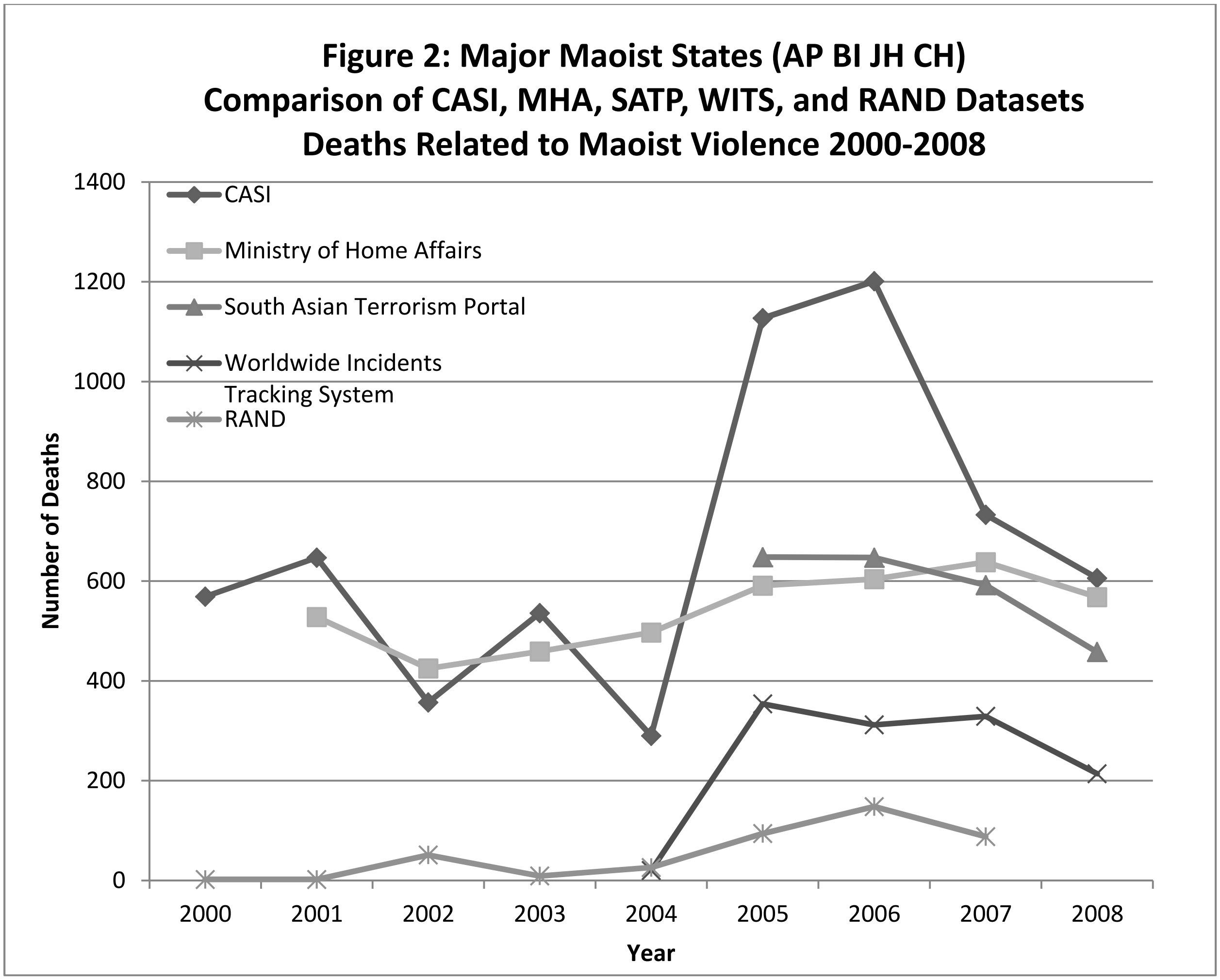


Figure 2: Major Maoist States (AP BI JH CH)
Comparison of CASI, MHA, SATP, WITS, and RAND Datasets
Deaths Related to Maoist Violence 2000-2008
1400
1200
1000
800
600
400
200
0
Number of Deaths
CASI
Ministry of Home Affairs
South Asian Terrorism Portal
Worldwide Incidents Tracking System
RAND
2000
2001
2002
2003
2004
2005
2006
2007
2008
Year

Table 1: Vegetation and Total Deaths from Maoist incidents
OLS and Negative Binomial (NB) models (Uninstrumented)

	Dependent Variable:				
	ln(Total Deaths)	Total Deaths	Δln(Total Deaths)	ln(Total Deaths)	Total Deaths
	OLS	NB	ECM	OLS	NB
ln(Total Deaths)$_{t-1}$	0.0802	0.0944	−0.920***		
	[0.0693]	[0.106]	[0.069]		
Vegetation$_t$	−13.99***	−21.87***		−13.87***	−21.75***
	[4.732]	[7.215]		[4.789]	[7.257]
Vegetation$_{t-1}$	−7.857*	−16.48**	−29.67***	−8.234*	−17.20**
	[4.565]	[7.399]	[10.01]	[4.569]	[7.287]
Vegetation$_{t-2}$	−7.827	−14.33*		−8.398	−15.10*
	[5.567]	[8.287]		[5.701]	[8.362]
ΔVegetation$_t$			−13.99***		
			[4.731]		
Δ Vegetation$_{t-1}$			7.827		
			[5.567]		
N	340	340	340	340	340
R^2	0.13	0.205	0.502	0.125	0.205
k	77	77	77	76	76
α		1.098***			1.105***
$z_{(\Sigma \text{ VEG}t)}$	−2.964***	−4.133***	-	−2.98***	−4.232***

Note:

1. Robust standard errors in brackets. Clustered (by disctrict) in OLS models.
   *** p<0.01, ** p<0.05, * p<0.1
2. All models include district fixed effects and year dummies. NB models estimated with true fixed-effects (in the model).
3. NB models: Dependent variable is Total Deaths. Pseudo R^2 reported.
   Error correction model (ECM): Dependent variable is Δ ln(Total Deaths). See Equation (2).
4. α is the overdispersion parameter in NB model. α > 0 indicates overdispersion;
   α = 0 indicates Poisson is appropriate.
5. All coefficients are to be interpreted as in a log-linear model.
6. $z_{(\Sigma\text{VEG}t)}$ tests the hypothesis: Vegetation$_t$ + Vegetation$_{t-1}$ + Vegetation$_{t-2}$ = 0.
7. Models without lagged dependent variables are reported to check for Nickell bias (Nickell 1981).

Table 2: First stage for IV results
Dependent variable: Lags of Vegetation Index (NDVI)

	Vegetation_t	$\text{Vegetation}_{t\text{-1}}$	$\text{Vegetation}_{t\text{-2}}$	Vegetation_t	$\text{Vegetation}_{t\text{-1}}$	$\text{Vegetation}_{t\text{-2}}$
$\ln(\text{Total Deaths})_{t-1}$	0.0002	–0.001	–0.001			
	[0.001]	[0.001]	[0.001]			
Rain_t	–0.104	–0.0144	–0.652***	–0.108	–0.0017	–0.639***
	[0.243]	[0.247]	[0.247]	[0.243]	[0.246]	[0.246]
Rain_{t-1}	1.393***	0.0828	–0.369*	1.396***	0.0709	–0.381*
	[0.211]	[0.238]	[0.216]	[0.211]	[0.239]	[0.217]
Rain_{t-2}	0.272	1.585***	0.0688	0.271	1.590***	0.0733
	[0.225]	[0.223]	[0.254]	[0.224]	[0.225]	[0.253]
Rain_{t-3}	1.189***	0.419	1.349***	1.181***	0.446*	1.376***
	[0.254]	[0.256]	[0.253]	[0.254]	[0.253]	[0.253]
District fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	340	340	340	340	340	340
R^2	0.218	0.222	0.206	0.217	0.219	0.202
k	78	78	78	77	77	77
Weak Instrument Diagnosis:						
Partial R^2	0.14	0.134	0.158	0.14	0.135	0.162
First-stage *F*	12.86	13.61	14.55	12,89	13.42	15
Kleibergen-Paap (WI)		10.19			10.25	

Note:

1. Robust standard errors clustered (by district). *** p<0.01, ** p<0.05, * p<0.1
2. Models include 68 district fixed effects and 5 year dummies.

Table 3: Vegetation and Total Deaths from Maoist incidents

IV models. *Dependent Variable* : Total Deaths (logged for OLS-IV)

	Vegetation				Vegetation Interactions		
	OLS-IV	**NB-IV**	**OLS-IV**	**NB-IV**	**NB-IV**	**NB-IV**	**NB-IV**
SECOND STAGE:							
ln(Total Deaths)$_{t-1}$	0.0398	0.104			0.070	0.0729	0.0793
	[0.0780]	[0.102]			[0.101]	[0.101]	[0.101]
Vegetation$_t$	−11.54	−15.05	−11.35	−14.75			
	[14.28]	[20.08]	[14.20]	[19.99]			
Vegetation$_{t-1}$	−8.831	−25.17	−8.905	−24.54			
	[13.31]	[22.47]	[13.31]	[22.24]			
Vegetation$_{t-2}$	−38.45***	−69.15***	−39.16***	−70.97***	−27.79		
	[14.49]	[23.60]	[14.38]	[23.76]	[39.76]		
Vegetation$_{t-2}$ × Veg$_t$					−22.29	−36.89	
					[35.33]	[33.27]	
Vegetation$_{t-2}$ × Veg$_{t-1}$					−66.71	−84.83**	−101.8***
					[43.23]	[33.70]	[37.00]
Vegetation$_t$ × Veg$_{t-1}$							−10.76
							[32.33]
N	340	340	340	340	340	340	340
k	77	77	76	76	80	79	80
α		1.107***		1.115***	1.094***	1.099***	1.103***
$z_{(\Sigma \text{ VEG}t)}$	−3.276***	−3.647***	−3.335***	−3.629***			
p-val $_{(\Sigma \text{ VEG}t)}$	0.001	0.0003	0.001	0.0003			
p-val $_{(\text{VEG}t - \text{VEG}t\text{-}1)}$	0.896	0.761	0.906	0.763			
p-val $_{(\text{VEG}t - \text{VEG}t\text{-}2)}$	0.251	0.081	0.230	0.071			
p-val $_{(\text{VEG}t\text{-}1 - \text{VEG}t\text{-}2)}$	0.173	0.245	0.167	0.219			
∂ ln(Total Deaths) /∂ Veg$_{t-2}$					−64.3***	−49.94***	−45.92***
					[25.46]	[11.80]	[11.12]
∂ ln(Total Deaths) /∂ Veg$_t$							−4.413
							[13.25]
FIRST STAGE:							
#Instruments	4		4				
Kleibergen-Paap (WI)	10.19		10.25				
Hansen's *J*	0.060		0.084				
Hansen's *J* (*p*- value)	0.807		0.772				
Anderson-Rubin (A-R)	3.547		3.752				
Anderson-Rubin (*p* val)	0.008		0.005				

Note:

1. Robust standard errors in brackets. Clustered (by district) in OLS models.*** p<0.01, ** p<0.05, * p<0.1
2. See Notes to Table 1
3. NB-IV models: Predicted endogenous variables in first stage used as regressors in the second stage.
4. ∂ ln(Deaths) /∂ Vegetation$_{t-2}$ reported at the means of other interactions.
5. p-val $_{(\text{VEG}t - \text{VEG}t\text{-}1)}$ is the p-value for the hypothesis test Vegetation$_t$ = Vegetation$_{t-1}$

Table 4: Robustness: Killings disaggregated by:
(1) # Civilians killed, (2) # Maoists killed, and (3) # Security Personnel killed

	OLS - IV				NB - IV		
	ln(Civilian)	ln(Maoist)	ln(Security)	ln(Sec)–ln(Mao)	Civilian	Maoist	Security
Vegetation_t	–14.83	–1.648	–17.84**	–16.63	–31.51	8.53	–54.78
	[10.65]	[12.06]	[8.655]	[14.05]	[26.19]	[25.52]	[38.51]
Vegetation_{t-1}	–14.25	–1.927	2.287	4.174	–53.31*	15.87	–25.39
	[11.02]	[11.01]	[8.497]	[11.36]	[29.61]	[27.58]	[34.42]
Vegetation_{t-2}	–22.79**	–22.65*	–12.77	9.682	–54.46**	–72.94**	–75.83**
	[9.868]	[12.45]	[8.758]	[14.39]	[26.88]	[30.62]	[36.74]
$\ln(\text{Civilian Deaths})_{t-1}$	–0.008				0.036		
	[0.094]				[0.150]		
$\ln(\text{Maoist Deaths})_{t-1}$		–0.016				–0.082	
		[0.075]				[0.122]	
$\ln(\text{Security Deaths})_{t-1}$			0.044				–0.128
			[0.103]				[0.198]
$\ln(\text{Total Deaths})_{t-1}$				–0.0177			
				[0.074]			
$z_{(\Sigma\ \text{VEG}t)}$	–3.652***	–1.712*	–2.494**	–0.147	–3.518***	–1.272	–3.301***
p-val $_{(\Sigma\ \text{VEG}t)}$	0.0003	0.088	0.013	0.883	0.0004	0.203	0.001
p-val $_{(\text{VEG}t - \text{VEG}t\text{-}2)}$	0.642	0.288	0.713	0.230	0.565	0.048	0.744
N	340	340	340	340	340	340	340
k	77	77	77	77	77	77	77
α					1.252***	1.332***	1.76***
#Instruments	4	4	4	4			
Kleibergen-Paap	10.4	10.12	10.22	10.19			
Hansen's J	5.475	1.451	1.628	0.127			
Hansen (p- val)	0.019	0.228	0.202	0.722			
Anderson-Rubin	4.575	1.632	2.881	0.485			
A-R (p-val)	0.001	0.167	0.023	0.746			

Note:

1. Robust standard errors in brackets. Clustered (by district) in OLS models. *** p<0.01, ** p<0.05, * p<0.1
2. See Notes to Table 3.

Table 5: Total Deaths from Maoist incidents and its association with Consumption Spending

Dependent Variable : ln(Total Deaths)

	Uninstrumented				
	OLS	NB	ECM	OLS	NB
ln(Total Deaths)$_{t-1}$	0.0897	0.103	−0.909***		
	[0.0627]	[0.105]	[0.062]		
Consumption$_t$	−0.613	−1.143		−0.635	−1.172
	[0.572]	[0.783]		[0.568]	[0.783]
Consumption$_{t-1}$	−0.917**	−2.160***	−1.515*	−0.947**	−2.213***
	[0.442]	[0.719]	[0.786]	[0.448]	[0.716]
Consumption$_{t-2}$	−0.00175	0.25		−0.0335	0.242
	[0.467]	[0.865]		[0.475]	[0.887]
ΔConsumption$_t$			−0.613		
			[0.578]		
Δ Consumption$_{t-1}$			0.073		
			[0.368]		
N	340	340	340	340	340
k	77	77	77	76	76
α		1.143***			1.152***
z_{CONSUMP}	−1.556	−2.003**	-	−1.584	−2.047**

	Instrumented	
	OLS-IV	NB-IV
ln(Total Deaths)$_{t-1}$	0.0465	0.104
	[0.109]	[0.102]
Consumption$_t$	−7.831	−12.45*
	[7.013]	[6.941]
Consumption$_{t-1}$	2.244	1.099
	[7.888]	[8.041]
Consumption$_{t-2}$	−8.859	−15.52**
	[6.419]	[6.839]
N	340	340
k	77	77
α		1.11***
z_{CONSUMP}	−2.023**	−3.771***
FIRST STAGE:		
#Instruments	4	
Kleibergen-Paap (WI)	0.496	
Hansen's *J*	0.054	
Hansen's *J* (*p*- value)	0.816	
Anderson-Rubin (A-R)	3.547	
Anderson-Rubin (*p* -val)	0.008	

Note:

1. Robust standard errors in brackets. Clustered (by disctrict) in OLS models. *** p<0.01, ** p<0.05, * p<0.1
2. See Notes to Table 1 and Table 3

Table 6: Controls: Spatial Effects, SC/ST, Mining, Inequality

Dependent Variable : Total Deaths (logged for OLS-IV)

	OLS-IV						NB-IV					
ln(Total Deaths)$_{t-1}$	0.0377	0.041	0.05	0.031	0.015	0.020	0.085	0.112	0.113	0.101	0.094	0.091
	[0.0765]	[0.078]	[0.078]	[0.078]	[0.080]	[0.078]	[0.099]	[0.103]	[0.101]	[0.103]	[0.100]	[0.097]
Vegetation$_t$	−8.955	−10.09	−12.74	−10.33	−11.11	−7.445	−9.682	−13.29	−17.54	−13.88	−15.21	−8.313
	[14.51]	[14.45]	[14.28]	[14.28]	[14.56]	[14.72]	[20.31]	[20.25]	[20.45]	[20.23]	[20.08]	[21.02]
Vegetation$_{t-1}$	−6.765	−7.461	−7.973	−8.333	−10.09	−3.724	−18.99	−22.04	−21.23	−24.58	−26.06	−4.906
	[13.13]	[13.64]	[13.35]	[13.46]	[13.10]	[13.28]	[22.24]	[22.95]	[23.02]	[22.95]	[22.27]	[23.55]
Vegetation$_{t-2}$	−35.49**	−39.57***	−37.45**	−40.60***	−41.86***	−40.65***	−67.13***	−70.01***	−69.17***	−71.51***	−73.11***	−74.03***
	[13.93]	[14.66]	[14.51]	[14.49]	[15.81]	[15.27]	[22.53]	[23.69]	[23.98]	[23.59]	[24.43]	[23.65]
Neighborhood2	0.264**					0.324**	0.326*					0.514***
	[0.125]					[0.131]	[0.167]					[0.178]
Proportion SC		0.916				1.022		1.443				2.338*
		[0.704]				[0.710]		[1.301]				[1.345]
Proportion ST			0.841			1.606**			2.101			3.333***
			[0.832]			[0.796]			[1.320]			[1.286]
Consumption GINI				1.076		1.409				0.952		1.412
				[1.162]		[1.181]				[1.467]		[1.476]
Value of Mining Output					0.059	0.069					0.082	0.105
					[0.048]	[0.048]					[0.087]	[0.097]
$z_{(\Sigma\ VEG t)}$	−2.74***	−3.17***	−3.25***	−3.29***	−3.35***	−2.69***	−3.10***	−3.50***	−3.58***	−3.67***	−3.73***	−2.73***
N	340	340	340	335	340	335	340	340	340	335	340	335
k	81	81	81	81	81	85	81	81	81	81	81	85
α							1.067	1.101	1.106	1.122	1.103	1.027
#Instruments	4	4	4	4	4	4	4	4	4	4	4	4
Kleibergen-Paap	10.26	9.76	10	10.35	9.684	9.409						
Hansen's J	0.019	0.037	0.056	0.075	0.016	0.003						
Hansen (p- val)	0.889	0.847	0.812	0.784	0.898	0.957						
Anderson-Rubin	2.743	3.519	3.453	3.789	3.576	2.845						
A-R (p -val)	0.029	0.008	0.009	0.005	0.007	0.025						

Note:

1. Robust standard errors in brackets. Clustered (by district) in OLS models. *** p<0.01, ** p<0.05, * p<0.1
2. See Notes to Table 3.

Table 7: Robustness: By State

Dependent Variable : Total Deaths (logged for OLS-IV)

	OLS-IV	NB-IV
$\ln(\text{Total Deaths})_{t-1}$	0.054	0.069
	[0.061]	[0.095]
Vegetation_t	−3.059	5.124
	[13.57]	[20.75]
Vegetation_{t-1}	−19.87	−42.60*
	[13.28]	[23.96]
Vegetation_{t-2}	−39.13**	−74.16***
	[14.98]	[25.84]
BIHAR × Vegetation_t	−7.416	−24.24
	[8.841]	[15.65]
BIHAR × Vegetation_{t-1}	9.27	16.77
	[7.699]	[15.99]
BIHAR × Vegetation_{t-2}	−7.709	−4.382
	[12.42]	[21.60]
CHATTIS × Vegetation_t	10.09	21.25
	[25.02]	[31.57]
CHATTIS × Vegetation_{t-1}	−4.477	28.92
	[18.77]	[24.02]
CHATTIS × Vegetation_{t-2}	44.87***	95.39***
	[8.697]	[19.31]
JHAR × Vegetation_t	−36.51**	−83.11***
	[17.77]	[28.72]
JHAR × Vegetation_{t-1}	−2.222	−2.885
	[12.36]	[23.21]
JHAR × Vegetation_{t-2}	6.546	−6.761
	[17.08]	[28.45]
$z_{\text{VEG (Andhra Pradesh)}}$	−3.641***	−3.497***
$z_{\text{VEG (BIHAR)}}$	−2.564***	−3.271***
$z_{\text{VEG (CHATTISGARH)}}$	−0.244	0.568
$z_{\text{VEG (JHARKHAND)}}$	−2.775***	−3.969***
N	340	340
α		0.98***

Note:

1. $z_{\text{VEG (Andhra Pradesh)}}$ is the z- statistic for the hypothesis of the *total effect* for AP: $\text{Vegetation}_t + \text{Vegetation}_{t-1} + \text{Vegetation}_{t-2} = 0$.

   $z_{\text{VEG (JHARKHAND)}}$ is the z- statistic for the total effect for Jharkhand:
   $(\text{Vegetation}_t + \text{Vegetation}_{t-1} + \text{Vegetation}_{t-2})$
   $+ (\text{JHAR}\times\text{Vegetation}_t + \text{JHAR}\times\text{Vegetation}_{t-1} + \text{JHAR}\times\text{Vegetation}_{t-2}) = 0$

Table A1. Media Sources for Data Base on Maoist Incidents

	State	Media Source
1	Andhra Pradesh	(E) Indian Express; The Hindu (V) Eenadu (WS) PTI; IANS
2	Bihar	(E) Indian Express; The Hindu; Times of India (Patna ed.) ; Telegraph (V) Hindustan; Prabhat Khabar (WS) PTI; IANS
3	Chhattisgarh	(E) Indian Express; The Hindu (V) Deshbandhu; Harit Pradesh; Navbharat; Hindustan (WS) PTI; IANS
6	Jharkhand	(E) Indian Express; The Hindu; Times of India(Patna ed.); Telegraph (V) Hindustan; Prabhat Khabar (WS) PTI; IANS

(E): English Language Daily
(V): Vernacular/Local Language Daily
(WS): Wire Services. PTI: Press Trust of India; IANS: India Abroad News Service (2000-2009).

Table A2: Descriptive Statistics

		Mean	sd	*N*
Total Deaths (#)	overall	11.51	45.44	340
	within		28.20	
Civilian Deaths (#)	overall	4.506	23.78	340
	within		17.72	
Maoist Deaths (#)	overall	4.353	13.27	340
	within		8.790	
Security Deaths (#)	overall	2.649	14.71	340
	within		9.828	
ln(Total Deaths)	overall	1.205	1.401	340
	within		0.844	
ln(Civilian Deaths)	overall	0.686	1.054	340
	within		0.711	
ln(Maoist Deaths)	overall	0.719	1.122	340
	within		0.701	
ln(Security Deaths)	overall	0.430	0.888	340
	within		0.605	
Vegetation (NDVI measure, range [−1, +1])	overall	0.411	0.310	340
	within		0.011	
Vegetation (Predicted)	overall	0.411	0.310	340
	within		0.005	
Consumption ('000 Rupees per month)	overall	0.640	0.243	340
	within		0.136	
Consumption (Predicted)	overall	0.640	0.223	340
	within		0.095	
Neighborhood2 (number of closest two	overall	0.903	0.802	340
districts with killings)	within		0.478	
Proportion Scheduled Caste (SC)	overall	0.187	0.118	340
	within		0.075	
Proportion Scheduled Tribe (ST)	overall	0.117	0.187	340
	within		0.059	
Consumption Gini	overall	0.257	0.068	335
	within		0.044	
Consumption 90 percentile / 10 percentile	overall	0.427	0.110	335
	within		0.076	
Value of Iron Ore Output (Rupees Billion)	overall	0.443	2.942	340
	within		1.532	
Value of Bauxite Output (Rupees Billion)	overall	0.013	0.060	340
	within		0.032	
Value of Mining Output (sum of iron ore and	overall	0.455	2.950	340
bauxite)	within		1.542	

Note: 2004-2008 for 68 districts in AP, BI, CH, JH.

This portion not for review

WEB APPENDIX 1: CASI DATABASE ON MAOIST INCIDENTS

Existing Data Sources

The official source on Maoist-related violence is released by the Ministry of Home Affairs of the Government of India. This data are divided into two categories: number of casualties (deaths, disaggregated by civilian, security forces and Maoist) and the number of violent incidents. However, these data are only available at the state level. Other popularly used data sources are the RAND-MITP Terrorism Incident database; data from the Worldwide Incidents Tracking System (WITS) from National Counter Terrorism Centre and data from the South Asian Terrorism Portal (SATP). The Rand-MIPT and the WITS data sets are world-wide and their Indian data is ad hoc and do a poor job at capturing Indian data reported by the non-English language press. The SATP data set has been assembled by the Institute for Conflict Management, available at their website, the South Asia Terrorism Portal. These data are based primarily on reports in the major English press, but do a considerably better job than the other two, especially in more recent years. However, this data set is only from 2005 and even after that it is not comprehensive.

The CASI database

Our Center for the Advanced Study of India (CASI) data set goes back to 2000. It has been compiled from multiple media sources. Even though two national English dailies – the *Indian Express* and *The Hindu* – have covered the Naxal issue extensively, they provide only partial coverage. We drew on ten other media sources: two additional English-languge newspapers: *Times of India* (Patna edition) and the *Telegraph*; six regional language press sources: *Eenadu*; *Hindustan*; *Prabhat Khabar*, *Deshbandhu*; *Harit Pradesh*; *Navbharat*; and two wire services: PTI and IANS.

We had to overcome a number of barriers to the creation of a reliable data set – access to available sources, cost of accessing data and the quality of data sources.

Access

Few newspapers have digitized archives that are freely available on the internet. Where available, we thoroughly documented that data. The remaining data were had to be acquired by contacting the various newspapers for access to their archives. We used a combination of methods. Our researchers contacted media offices directly. Since Maoist violence is a politically extremely sensitive topic locally, some media offices were understandably apprehensive about our request and others cited prohibitively high user fees. To overcome this hurdle, we fell back on the networks between the Indian academic community and senior, influential media persons to be able to secure access. Regional newspapers were more forthcoming and prompt in granting access as opposed to national dailies (both English and Hindi).

Another problem faced by researchers was where newspapers had handed over the management of their archives to private companies specializing in maintaining the documents of large companies. In that case, the agent company decided the terms and conditions of access, which usually came at prohibitive cost. We were able to obtain access only after intervention by editors and some protracted discussions between researchers, the media house and the company where the archives were stored.

Cost

Some media houses charge a daily user fee, apart from fees for photocopying the relevant material. The cost of accessing archives managed by private documentation companies is very high. Fees are charged under various heads – a daily user fee, a fee per issue of newspaper viewed, a fee per monthly file viewed, and 20 per cent additional charge in lieu of the storage charges being paid by the media house to the private company. Suffice it to say that the data and field work come at a significant cost. We believe the data to be far superior to the sources mentioned in the first paragraph.

Quality

In most cases newspapers were stored in poor conditions, with issues for an entire month often bound into a single file for the purpose of storage. In most of the archives, the files were not stored in any systematic manner – for instance year-wise – that could ensure easy access. A lot of effort was spent re-ordering the files. Since files were sometimes missing relying on a single source for local data would compromise the quality of the data. We have therefore followed the more costly strategy of scouring multiple sources for each state. We therefore have attempted to ensure as much overlapping coverage as possible, making the data far more reliable than any other data source.

Uneven Reporting

Apart from access, reportage of incidents is uneven. The fact that different incidents were reported on the same day in different newspapers implies that several incidents are not being reported, and hence, considerable underreporting of the problem by any individual newspaper. Coverage in national newspapers is stronger in some regions than others. For some years the overlapping coverage of incidents in two national English-language newspapers – The Hindu and Indian Express – was less than a fifth, while the two combined were a small fraction of the total number of incidents in the final CASI database, since it was built from multiple sources.

Coding

For each incident we have intended to record the following:

- date (month and year)
- location (district and village)
- type of incidence (bomb explosion, kidnapping etc…)
- number of Maoists killed in that incident
- number of non-Maoist civilians killed
- number of security personnel killed
- other information relevant to the incident

While different members of our team accessed news sources from different parts of the country, a copy of each story was sent to a core team for coding to ensure consistency. The exception was *Eenadu* which needed a person fluent in Telugu as well as with a strong social science background so that the coding could be done directly.

Even this painstaking effort is not without problems. When several newspapers report the same incident, the news is sometimes printed on different dates. News reports often carry different placelines (place of reportage), as well as different accounts of the location of the incident. Details on the number of casualties and injuries, how many of those involved were Maoist, how many were civilians and how many were security forces have different figures in different accounts. In some stories, the number of injured and dead remains unspecified. Unless an incident of Maoist violence is large in scale or intensity, it is rarely followed up.

A frequently used phrase during reportage is 'suspected Maoists' and 'Maoist sympathizers'. In doubtful cases, where during the time of reportage, the journalist is unsure about details and the discrepancies are rarely resolved through follow-up articles. We have coded them as civilians.

Another issue relates to the *Salwa Judum*, a militia consisting of local tribal youth organized by the government of Chhattisgarh to counter the Naxalite violence in the region. Since members of this militia had the status of Special Police Officers (SPOs) and received training and wages from the Chhattisgarh state government we have coded them as security forces (A judgment of the Indian Supreme Court in 2011 declared this militia as illegal).

Our intention in providing these details is to indicate that data pertaining to Maoist casualties is not problem-free, but we have taken as much care as possible to present the most representative picture to date. All politics is local, and so is all political violence. The Rand-MIPT and the WITS data are selected data from large newspapers, and considerably underreport the actual extent of the conflict. When an event is newsworthy, it is better covered than if it is confined locally. For an issue with enormous policy economic and political implications as the Maoist conflict, we think the data must be up to the task. We think ours is.