



Working Paper 12-18  
Economic Series  
June, 2012

Departamento de Economía  
Universidad Carlos III de Madrid  
Calle Madrid, 126  
28903 Getafe (Spain)  
Fax (34) 916249875

## The Political Economy of the Maoist Conflict in India: An Empirical Analysis

First draft- March 2011, This draft- June 2012

Joseph Flavian Gomes\*  
Universidad Carlos III de Madrid

### Abstract

This paper contributes to a burgeoning literature that uses sub-national micro data to identify the causes of civil conflicts. In particular, we study the Maoist/Naxalite conflict in India by constructing a comprehensive district level database using conflict data from four different terrorism databases and combining it with socioeconomic and geography data from myriad sources. In addition to exploiting the within country regional heterogeneity, we use the micro structure of the data to construct group-level characteristics. Using data on 360 districts for 3 time periods, we find evidence on how land inequality and lower incomes are important for the Maoist conflict. Moreover, making use of the micro structure of the data we are able to ask whether exclusion of the low castes and tribes from the growth story of India is important. We find that while the income levels of the different ethnic groups are not important, the growth of incomes of Scheduled Tribes significantly decreases the intensity of the conflict. Finally, we show how historical property rights institutions from colonial times that go back centuries can affect present day conflict outcomes through their impact on economic outcomes, social relations and the political environment in the district.

\*PhD candidate, Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, Getafe 28903, Madrid. E-mail: jgomes@eco.uc3m.es. I would like to thank Klaus Desmet, Irma Clots, Ulrich Wagner, Raquel Carrasco, Diego Puga, Juan Dolado, Abhiroop Mukhopadhyay and Rohini Somanathan. I am grateful to all seminar participants in the UC3M PhD student workshop, UC3M SOS seminar, ECORE summer school 2011, NCDE-Nordic Conference on Development Economics 2011; CRED Workshop, CRED, (FUNDP), Namur, 2011; LICOS seminar, LICOS Centre for Institutions and Economic Performance, KUL, Leuven, 2011; for their comments and feedback.

# 1 Introduction

In recent years the relation between economic performance and civil conflicts has generated a considerable amount of interest among economists. Not surprisingly, within the span of a few years a lot has been written on the subject.<sup>1</sup> A good part of the literature has taken a cross-country approach using aggregate data to identify the causes of civil conflicts.<sup>2</sup> However, there is a small but burgeoning literature showing that going to the sub-national level is key. Conflicts are often localized, and have to do with the unequal spatial distribution of resources within countries. Thus, treating countries as being homogeneous is often problematic. In addition, the use of aggregate data in most studies limits the kind of questions one can address. For example, if ethno-linguistic diversity is shown to have an impact on the probability of conflict, aggregate data do not allow us to determine whether this is solely picking up the effect of cultural diversity or whether it is also proxying for economic heterogeneity across groups. To address this question, one needs microdata that give information on, say, income at the level of different ethnic groups.

This paper uses micro data at the sub-national level to analyze the Maoist (aka Naxalite) conflict in India. The conflict started as a localized land conflict in Naxalbari (hence the name Naxalite), a village in West Bengal in 1967. However, it has seen a terrifying increase in proportions only in the last decade. In the period 2004-2010 there have been more than 5000 lives lost (even by official estimates). Including the number of wounded and displaced would make the figure many times higher. In fact, it has been identified as “the single biggest security challenge to the Indian state” by Dr. Manmohan Singh, the Prime Minister of India. While on paper the aim of the movement is to establish a “people’s democratic state under the leadership of the proletariat” (Harris (2010)), at the heart of the conflict is land (rights, acquisition and its unequal distribution) and “in practice land redistribution appears to be one of the main goals” (Iyer (2009)). In addition, there is an ethnic/caste element to the conflict, as some tribal groups are at the lower end of the income distribution, and feel they are being left behind the rising tide of the Indian economy in the last decades (Guha (2007)).<sup>3</sup>

Although the conflict has spread over several states across India, by no means is it affecting

---

<sup>1</sup>See Blattman and Miguel (2010) for a recent survey of the existing literature.

<sup>2</sup>See for e.g. Collier and Hoeffler (2004), Fearon and Laitin (2003), Miguel et al. (2004), Ciccone (2010), Besley and Persson (2008).

<sup>3</sup>Moreover, a lot of the land acquisition is taking place in the tribal areas.

all regions in the same way. The goal of the paper is to exploit this spatial heterogeneity to understand the sources of the conflict. In particular we try to address some of the following questions: How important is land inequality? Are tribal groups resorting to violence because of being left behind? Does the spatial heterogeneity in colonial institutions help us explain the current distribution of violence?

With the above goal in mind we use district level conflict data for the period 1979-2009 along with socio-economic and geography data from multiple National Sample Survey (NSS) sample surveys, Censuses etc. to build a district level dataset. To account for differences across tribal groups and castes, we use microdata to construct economic variables at the level of these groups for each district. This gives us a comprehensive dataset of 360 districts (for the 16 main states which constitutes  $\succ 90\%$  of the population) over three time periods. We use Probit regressions to explain the probability of conflict and Negative Binomial regressions to explain the intensity of conflict at the district level. The main findings of the paper are listed below.

A first finding is that land inequality is one of the key determinants of the conflict controlling for all other factors that the literature has found to be important. Land inequality reflects not just the inequality in the distribution of land but also differences in the socio-economic lives of people in a predominantly agrarian society. Moreover, it also implies more scope for inadequate compensation under land acquisition.

District income on the other hand, also comes out to be an important determinant of the conflict. In the Probit and Negative Binomial regressions low income comes out to be a significant predictor of conflict. Moreover, we obtain similar results in the 2SLS, Instrumental variable regressions using its own lagged values as an instrument for income. We find that lower income significantly increases both the probability and the intensity of conflict.

The evidence on the being-left-behind hypothesis of the disadvantaged groups, i.e., the low castes and the tribes, is mixed. One often heard argument is that tribal groups and lower castes resort to violence because their groups are not equally benefiting from the high rates of growth. We find that the income levels of the Scheduled Castes or Scheduled Tribes have no significant impact on the conflict. However, when we look at the growth in incomes of the three groups separately, the picture changes. We find that a lower growth in incomes of the Scheduled Tribes significantly increases the intensity of conflict. Moreover, in some specifications, we also find that the presence of the Scheduled Tribes leads to more conflict.

A final finding is that historical institutions matter. Class antagonism driven by land institutions that have lingered for centuries has a significant impact on both conflict presence and intensity. Districts where land rights were traditionally enjoyed by landlords have higher conflict compared to districts where land rights were traditionally with the farmers themselves.

The existing literature has witnessed several different approaches to empirically identify the causes of civil conflicts. There are two clear directions in which this literature needs progress. The first direction is using sub-national micro data in order to overcome the shortcomings of the cross country analyses. The second crucial issue is to establish a causal relation between conflict and its determinants.

With regard to the use of cross country data in the analysis of civil conflict, Do and Iyer (2009) point out two caveats: (1) Data might not be comparable across countries. (2) Reasons for the conflict might vary from country to country.<sup>4</sup> Moreover, another serious shortcoming of such studies is that they ignore the within country heterogeneity by treating the country as a unit of observation.<sup>5</sup> However, conflicts are often localized and depend on the unequal spatial distribution of resources within the country. For example, in the context of the Maoist conflict in India, in West Bengal, one of the severely affected states, the conflict is very pronounced in the Midnapore and Puruliya districts while it is completely absent in districts like Howrah, North and South 24 Parganas.<sup>6</sup> This is the kind of heterogeneity that one cannot take into account using even states as units of analysis. Thus, using countries as the units of analysis would be unsatisfactory.

In addition, “the country level of analysis has [other] inherent limitations. Individual and group level conflict factors such as poverty and ethnic hostility are imperfectly tested at the national level” (Blattman and Miguel (2010), Sambanis (2004)). By making use of micro data we can take into account the differences in incomes of disadvantaged groups vis-a-vis others and also the heterogeneity in the distribution of these groups. Moreover, the sudden rise in the conflict levels from 2005 onwards is surprising given that the country has been experiencing above 8% growth

---

<sup>4</sup>Further, even within a country focusing on one specific conflict might give us more interesting insights than looking at overall violence.

<sup>5</sup>Also, they do not allow us to control for all the factors that are constant within the country viz. the macroeconomic variables.

<sup>6</sup>Again, if one looks at Maharashtra (which is a state located in the west of the country), almost all the Maoist incidents have been concentrated solely in the Gadchiroli district (formerly part of the Chandrapur district). Again, while the conflict affects virtually all the districts of Bihar, it is almost completely absent in the states of Rajasthan, Punjab and Haryana.

rates in the same period. The use of micro data could be particularly helpful in identifying regional and group level heterogeneities in income and growth rates of income that might be leading to more conflict. As Blattman and Miguel (2010) put it “... the most promising avenue for new empirical research is on the sub-national scale, analyzing conflict causes, conduct, and consequences at the level of armed groups, communities and individuals”.

Part of the lack of sub-national level studies has of course been due to the unavailability of reliable sub-national data and/or the tedious task of building such datasets. However, recent times have witnessed a growing number of studies that seek the causes of civil conflicts using sub-national micro data with some interesting insights. Do and Iyer (2009) for example using a district level dataset on the Maoist Conflict in Nepal, show how the lack of economic opportunities triggers conflicts while suitable geographic conditions help it to perpetuate. However, ethnic and caste polarization, land inequality and political participation don't affect violence. Dube and Vargas (2008) show how positive shocks to the agricultural sector reduce conflict while positive shocks to the resource sector increases conflict in Colombia. Jha (2008) in a town level study of India shows how historical institutions determine peaceful coexistence or conflict between Hindus and Muslims. Iyer (2009), through district level regressions on India and Nepal, shows how lagging districts experience more conflict than others. From the above examples it is evident how micro level studies are a clear advance over the cross country studies.

The other critical issue in this literature is establishing a causal relation between conflict and its determinants. This is due to two main problems, as mentioned by Do and Iyer (2009) (1) There might exist unmeasured factors that affect both conflict intensity and pre conflict characteristics. (2) Districts that are experiencing more violence might also be districts that have experienced high past conflict. Recent studies have tried to address this issue using an instrumental variable approach when clearly exogenous instruments have been available. Miguel et al. (2004) and Ciccone (2010) have used weather shocks. Dube and Vargas (2008) have used exogenous shocks to agricultural and resource prices. When clear instruments have not been available authors have tried to use data on covariates from the pre conflict period in order to prevent endogeneity arising out of reverse causality.<sup>7</sup> Also, all potential covariates are controlled for in order to reduce endogeneity arising out of omitted variables. Following in the same vein, in this paper we use data from the pre conflict period; control for all possible variables (that the literature suggests are important) that might play a role subject to the data availability and always control

---

<sup>7</sup>e.g. Do and Iyer (2009); Mitra and Ray (2010)

for the presence of past conflict. In addition, we also run IV regressions with previous period income as an instrument for present income.

The other contribution of this paper is that it traces the origins of civil conflict to a new channel over and above the channels like underdevelopment and geography already identified in the literature. We see how historical colonial institutions from centuries ago can impact conflict outcomes in the present times. Banerjee and Iyer (2005) show that the colonial land revenue institutions set up by the British in India, lead to sustained differences in economic outcomes in the post independence period. Moreover, they observe that “the class based antagonism that [the land revenue institutions] created within the communities in these areas has persisted well into the post independence period.” “Areas most associated with Maoist uprisings are WB [West Bengal], Bihar & Srikakulam district of AP [Andhra Pradesh]- all landlord areas. Paul R. Brass (1994 pg. 326-327) argues explicitly that these peasant movements have their roots in the history of exploitation and oppression of peasants by landlords.” (Banerjee and Iyer (2005)) In this paper we empirically verify that these institutions indeed have an impact on the conflict.

As far as the literature on the Naxalite conflict itself is concerned, there are very few rigorous empirical studies. Barooah (2008) relying on a simple cross section OLS analysis finds that the probability of conflict in the district is increasing in the poverty rate and is decreasing in the literacy rate. Hoelscher et al. (2011) also using a cross section and relying on probit and negative binomial techniques, find forest cover, prevalence of conflict in the neighbouring district and presence of Scheduled Castes and Tribes to be important. Gawande et al. (2012) using a district level panel find that negative natural resource shocks increase the intensity of conflict. Vanden Eynde (2011) on the other hand also using a district level panel finds that negative labour income shocks increases violence against civilians to prevent them from being recruited as police informers. While all these papers are important for understanding the nature and causes of the Maoist conflict in India, they ignore some important factors like land inequality, historical land institutions and the exclusion of the tribals in India, which are crucial in understanding the conflict.

Thus, to summarize, this paper contributes to several different strands of the literature. The first strand is the research using sub-national-micro data exploiting the spatial heterogeneity within a country and the micro characteristics of the data to pin down the causes of civil conflict. This is a clear progress over existing cross country literature. Moreover, we show how horizontal inequality in growth rates matters rather than growth itself. We see that while overall growth/or

the lack of it does not affect the conflict, the low growth in incomes of the Scheduled Tribes significantly increases the intensity of conflict, controlling for income growth of other ethnic groups. Finally this paper also contributes to the broad class of literature that traces divergences in current economic outcomes to differences in historical institutions in a country.<sup>8</sup> We show that, in addition to economic underdevelopment, land relations and historical institutions within a country could lead to conflict.

In the next section we discuss the Maoist conflict in India, in Section 3 we list the main hypotheses of the study. Section 4 gives the empirical analysis and results and Section 5 concludes.

## 2 The Maoist/Naxalite Conflict

The start of the Maoist conflict is marked by a peasant uprising in the year 1967 in Naxalbari, a small village in the Darjeeling district of West Bengal. “A tribal youth having obtained a judicial order went to plough his land on 2 March 1967. The local landlords attacked him with the help of their goons. Tribal [peasants] of that area retaliated and started forcefully recapturing their land” (Kujur (2008)). The rebellion left nine tribal people and one police personnel dead and the Naxalite movement in India was born. The rebellion itself was contained by government forces within 72 days with the use of force, but had already gathered huge visibility from Communist revolutionaries from across the country.

After West Bengal the movement spread to the state of Andhra Pradesh where the formation of the People’s War Group (PWG) in 1980 marks the revival of the movement post the Naxalbari uprising. It has since then spread across various states in India including Bihar, Jharkhand, Madhya Pradesh, Orissa, Chhattisgarh, Maharashtra and Karnataka across many districts and has existed in varying degrees across the country.<sup>9</sup> However, it was the 2004 merger of the PWG with the Maoist Communist Center (MCC) that led to the formation of the Communist Party of India-Maoist (CPI-Maoist) that marks the modern revival of the movement and followed a huge rise in insurgency and violence thereafter.

While the term “Naxalite” comes from the place of birth of the movement the term “Maoist” is used due to the Maoist ideologies that many of these rebel groups adhere to. The CPI-Maoists for example claim to be committed to a “democratic revolution” through “a protracted

---

<sup>8</sup>A very novel concept in the conflict literature. Jha (2008) is the other paper in this context.

<sup>9</sup>See map in Figure B.1 in appendix B.

people's war with the armed seizure of power remaining as its central and principal task" (SATP; Iyer (2009)). While the conflict as a whole has been termed as a Maoist/Naxalite conflict, the movement itself is hardly homogeneous. It has in fact, always had a fragmented structure with multiple groups operating without a centralized movement organization. At present, the CPI-(Maoist) is the biggest operating group.

While in some small pockets of India like the Dantewada district in Chattisgarh the movement has reached a war-like situation with terror and violence on a daily basis, the entire movement has not always nor everywhere been violent. Some of the non-violent actions have included meetings, closures, economic and social boycotts, people's courts, rallies, road blocks, effigy burning, strikes etc. Even songs and plays have been used to mobilize people.<sup>10</sup> However, at least in the last five years the violent part of the movement has taken precedence. On April 6th, 2010, for example, several hundred Maoist guerrillas attacked a convoy in a forest in Chhattisgarh, killing 76 armed policemen (The Economist). In 2011 alone, by March 13, already 116 people had been killed due to the conflict (SATP).

There have been more than 5000 lives lost (civilians, rebels and security personnel) and more than 12000 incidents of violence in the period of 2004-2010 due to the Naxalite conflict. The number of people displaced in Chhattisgarh alone was more than 43000 (Sundar (2008)), end of 2006. The geographical spread of the conflict has seen a phenomenal increase from 55 districts in 8 states in 2003 to 194 districts spread across 18 states in 2007 (SATP; Iyer (2009)). "While it is difficult to put an exact figure to the number of rebels, there is an estimated 10000 to 20000 full time fighters with countless thousands of village militias controlling particularly remote jungle areas where the state is hardly present." (The Economist).

Surprisingly however, there are very few in-depth empirical analyses of the Maoist conflict. The few studies that exist<sup>11</sup> give us some interesting insights on the conflict, but at the same time they ignore several key issues. Thus, the conflict is perhaps best understood through the various case studies that exist on the issue. Harris (2010) gives a summary of the recent writings on the Naxalite conflict. Bhatia (2005) in an analysis of the conflict in Central Bihar, provides an exhaustive picture of the conflict. The Government of India has also produced descriptive studies on the issue.

---

<sup>10</sup>See Bhatia (2005)

<sup>11</sup>Barooah (2008), Hoelscher et al. (2011) Vanden Eynde (2011) Gawande et al. (2012)

In the existing case studies, several different factors have been identified as potential causes of the conflict. Bhatia (2005) for example, focusing on Central Bihar, identifies three distinct albeit inter-related classes of reasons behind the Naxalite movement viz. (1) Economic rights (2) Social rights (3) Political rights. The economic issues that Bhatia (2005) mentions are, land rights; minimum wages; common property resources; and housing. The Government of India in its own study recognizes land related factors, displacement and forced evictions, livelihood, and social oppression as the main socio economic reasons behind the discontent of people and support to Naxalism (Government of India (2008)).<sup>12</sup> Guha (2007) on the other hand argues that the exclusion of the tribals from the development in India is crucial to the conflict.

We focus on three of the key issues that seem to be recurrent as much in the literature as in the popular discussion about the conflict. The unequal distribution of land and the apparent exclusion of the Scheduled Tribes from the economic success of India are two of the key issues. Moreover, the overall lagging behind of some regions in India is the other important issue.

Land rights are indeed one of the most important issues taken up by the Naxalites: “*Khet par adhikar ke liye ladho, desh mai janwad ke liye badho*” (Fight for land rights, march towards democracy in the country - Liberation (a Naxalite group) slogan) (Bhatia, 2005). In fact, regardless of the ideologies of the different Maoist factions in practice land redistribution has remained one of the main goals of the movement. This is evident from the failed peace talks between the Andhra Pradesh government and the PWG in 2004 where this was one of the main issues (Iyer, 2009). Indeed the focus of the Naxalite movement is on trying to provide land, whether the land of landlords or government land, to the landless. In fact, as mentioned above the ‘Naxalite’ term is itself derived from the peasant uprising arising out of land disputes in Naxalbari. However, there are instances of such struggle even later and elsewhere in the country. “[For example in Bhojpur, Bihar,] Jagdish Mahato, a local teacher who had forged links with Naxalite leaders from West Bengal, led a protracted struggle against exploitative landlords” (Bhatia (2005)).

While the inequality in land distribution and the strife between big landowners and landless peasants seem to very important in Bihar and West Bengal, in Chhattisgarh the strife between poor tribal landowners and the state seems to be more important. In the recent past the government has been trying to acquire land for different purposes including dams, mining and

---

<sup>12</sup>Apart from the socio economic reasons issues arising out of bad governance and policing are also mentioned.

private industry. The Naxalites have been vehemently opposing the land acquisition by the state. “People are fighting against land acquisitions and the government is labeling them Naxals and then using that to suppress them. .... If the Naxals were not there, the government would be able to acquire more land for the private sector” (Sundar (2008)).

Underdevelopment has also been seen as one of the key reasons behind the rebellion. This is not surprising given that the movement is concentrated in some of the most impoverished regions of the country. “Even if there was no truth behind the accusations that poor tribes are being exploited and are being illegally disposed of their lands, it is fairly certain that the Naxalites are feeding on the festering discontent of the impoverished and marginalized tribal communities. According to the 2001 census for example, about three quarters of Dantewada’s 1,220 villages are almost wholly tribal; 1,161 have no medical facilities; 214 have no primary school; the literacy rate is 29% for men and 14% for women.” (The Economist, 2006).

As far as the participation is concerned, “... the social base of the movement ... consists overwhelmingly of the landless, small peasants with marginal landholdings, and to lesser extent middle peasants. In caste terms, the base of the movement consists of lower and intermediate castes” (Bhatia, 2005). The fight against the social oppression that the dalits [lower castes] and the lower among the OBCs [other backward castes] have been regularly subjected to is perhaps the most significant among the issues used by the Naxalite movement” (GOI, 2008). Moreover, there is a huge tribal participation in the movement. The Scheduled Tribes are in fact economically one of the worst performing groups in India (even behind the Scheduled Castes) and exclusion of the tribes from the growth that the Indian mainstream has been experiencing is seen as one of the key driving forces of the Maoist movement.<sup>13</sup>

Given the above discussion on the Naxalite conflict we now come to the main hypotheses of the study.

### **3 The Main Hypotheses**

The literature speaks of a variety of different factors that might lead to civil conflicts. “Civil Wars are more likely to occur in countries that are poor, are subject to negative income shocks, have weak state institutions, have sparsely populated peripheral regions and possess mountainous terrain” (Blattman and Miguel (2010)). Some of the other factors that have been mentioned

---

<sup>13</sup>See Guha (2007)

in the literature are, ethnic and religious diversity and fragmentation, lack of democracy and civil liberties, linguistic and religious discrimination, inequality, new states and political instability, non-contiguous territory, population pressure, colonial occupation, etc.<sup>14</sup>

In this study we try to identify which factors might be more relevant in the context of the Naxalite conflict in India. Thus, combining the findings from the previous literature and our understanding of the conflict we test for different hypotheses.

**Hypothesis 1: Land inequality increases conflict.**

Land inequality and exploitation of the landless/poor farmers by the wealthy landlords has been attributed to be one of the most important causes of the Naxalite movement. In fact, that's how it started in the first place. In a predominantly agrarian economy, the importance of land cannot be overemphasized and survival of the poor often depends on their access to land. There is in fact empirical evidence on the importance of land distribution on conflicts from elsewhere in the world as well. Andre and Platteau (1996) study a highly densely populated area in the Northwest of Rwanda during the period 1988-93 and find how the land distribution had become increasingly unequal and land dispossession rampant. They further show that pervasive incidence of land disputes and the threat of landlessness had led to rising tensions in social relations. They relate these adverse conditions in land distribution to the civil conflict that broke out in 1994. Verwimp (2003) offers an empirical analysis of the peasant participation in the Rwandan genocide. Using data from pre-genocide agricultural survey sample he builds a panel dataset. He finds that the genocide perpetrators comprised mostly of poor wage workers and land renters, while the victims were primarily from the landlord class. Land inequality has also been found to be important in the context of the Maoist conflict in Nepal. Macours (2011) shows how a disproportionate increase in the returns to land for the households with relatively large holdings and the corresponding relative deprivation of the (near) landless between 1995 and 2003, led to expansion of Maoist recruitment activities beyond their initial heartlands.

There are several potential ways in which the land distribution could affect the conflict. A highly skewed land distribution also reflects higher disparities in the social and economic lives of the people and thus a higher potential for grievance. Moreover, if the distribution is too unequal and dominated by a few large landlords while the vast majority of the population being small and landless then there is an additional source of problem. In a growing India the government is

---

<sup>14</sup>See Fearon and Laitin (2003) for a discussion.

trying to acquire land for a variety of purposes including industrialization, mining and building dams. In case of land acquisition by the government while the entire community is adversely affected the compensation mostly or fully goes to the big landlords who to start with enjoy a higher socio-economic status, while the vast majority of the affected are left uncompensated. Moreover, it is also common that tribal people in remote places have been living on a certain plot of land for generations but do not have any formal title deeds to the land resulting in no compensation whatsoever in case of acquisition by the government. Thus, the Naxalites are fighting for land issues against both big landlords and the state.<sup>15</sup>

The land distribution is crucial not only for the Maoist conflict but for the overall health of the rural economy. The land reforms in India in the post independence period thus sought to directly improve the access to land of the poor households. Besley and Burgess (2000), in a state level analysis of the Indian land reforms find that indeed the land reforms have been associated with significant poverty reduction. However, implementation was hardly homogeneous across states. These heterogeneities exist even within states which have witnessed widespread implementation. “The 3 extreme Maoist districts in West Bengal are West Midnapur, Puruliya, Bankura-where land reforms on the scale affected by the CPI-(M) in other parts of the state haven’t taken place.” (Chakravarti, (2008)) Thus, we test to see if indeed more land inequality is associated with more conflict.

### **Hypothesis 2: Underdevelopment leads to more conflict.**

The relation between income and civil conflict is one of the most robust relations in the empirical literature on conflicts. Collier and Hoeffler (1998, 2004) and Fearon and Laitin (2003) show that lower per capita incomes lead to higher probabilities of conflict. Collier and Hoeffler (2004) argues that the opportunity cost of fighting is lower when incomes are low and thus making it easier to recruit rebels. Fearon and Laitin (2003) on the other hand argues that lower incomes reflect limited state capacity to put the rebellion down. In other words, low per capita income is an indicator of weak and corrupt governments who might be unable to devise effective counter insurgency policies.

Apart from the levels of income, negative shocks to income have also been found to be important. The theoretical idea goes back to Becker (1968) where he argues that rise in returns to crime

---

<sup>15</sup>In a recent hostage incident when a government collector was kidnapped, two of the primary demands of the Maoists were as follows: grant of land rights to tribal people in scheduled areas; minimum displacement of tribals while making space for industries and mining (The Hindu, Feb 22, 2011: Orissa accepts eight Maoist demands: <http://www.hindu.com/2011/02/22/stories/2011022264251300.htm>).

induces more workers to the criminal sector. Miguel et al. (2004) demonstrate how negative income shocks lead to higher levels of conflict. Dube and Vargas (2008) argue that lower wages increase conflict (lower international coffee prices lead to a negative shock to certain regions of Colombia leading to higher conflict levels).<sup>16</sup>

Thus, a priori both income levels and changes in incomes could potentially affect the conflict. Districts with lower per capita incomes are expected to experience more conflict either through the opportunity cost channel or since government machinery is also more underdeveloped in poorer regions. Lower incomes could also generate grievances among the people fueling conflict. Moreover, we also run some regressions exclusively to test whether the heterogeneous growth rates across districts plays a role in the conflict.

**Hypothesis 3: The exclusion of disadvantaged groups (Scheduled Castes (SC) and Scheduled Tribes (ST)) leads to more conflict.**

In India, Scheduled Castes and the Scheduled Tribes are two groups that have been historically disadvantaged and find themselves at the bottom of the social hierarchy.<sup>17</sup> The Scheduled Castes are at the bottom of the Hindu caste system. The Scheduled Tribes on the other hand are tribal populations who are outside the Hindu caste system. As discussed in the previous section the main support and the recruits for the conflict are supposed to be mostly from the lower caste and tribal populations.<sup>18</sup> While these disadvantaged minorities have enjoyed affirmative action in the post independence India, they are still not economically and socially at par with others. “There are over 200,000 pending cases of atrocities against lower castes in India, and the conviction rate is just a little over 2 percent. There are an estimated 162 million untouchables in rural India, according to the National federation of Dalit Land rights movement. 70% of them don’t own land” (Chakravarti (2008)). Guha (2007) stresses that in the last 60 years after independence the tribals have been the group that has been left out and even suffered in the growth story of India. In fact, the predominantly tribal areas like Chhattisgarh have experienced very high levels of conflict.

Thus, caste and tribal identities are important issues that could exacerbate the conflict. It

---

<sup>16</sup>This is the opportunity cost channel. Dube and Vargas (2008) also find that a positive shock to the natural resource price increases conflict via the rapacity channel.

<sup>17</sup>“Historically disadvantaged groups are commonly defined as groups which have been systematically excluded from institutions and cultural practices that provide skills and resources.” Pande (2003)

<sup>18</sup>“The main support for the Naxalite movement comes from dalits (SC) and adivasis (ST) (Government of India (2008)).”

would be interesting to see once we control for incomes, if presence of lower castes and tribals still makes a difference. Moreover, heterogeneity in incomes across these ethnic lines could also be important. “Some case studies [Sambanis 2005, Frances Stewart 2001] suggest that ‘horizontal’ inequality- inequality that coincides with ethnic or other politically salient cleavages is a particular important driver of civil conflict.” (Blattman and Miguel (2010)) Thus, we try to verify whether grievances arising out of feelings of exclusion of the disadvantaged groups impact the conflict.

#### **Hypothesis 4: Historical land institutions directly impact the conflict.**

There is a huge body of literature arguing how historical institutions have a persistent effect on current economic performance.<sup>19</sup> In the context of the civil conflict literature, Jha (2008) has an interesting finding in his paper on the Hindu-Muslim communal conflicts in India. His primary conclusion is that, in the presence of a historical complementarity between the two groups the probability of conflict is lower. Besley and Persson (2008) on the other hand highlight the importance of political institutions in explaining conflict outcomes. While their main conclusion is that higher market prices of exported and imported primary commodities have a positively significant effect on the incidence of civil conflict, they find that those effects are heterogeneous depending on the quality of political institutions in the country.<sup>20</sup>

In the context of India, Banerjee and Iyer (2005) have shown that districts in which property rights in land were historically given to landlords by the British have ended up with worse economic outcomes in the post independence period. These areas have significantly lower investments in agriculture, health, education and also lower agricultural productivity than areas in which these rights were given to the cultivators. These historical institutions could affect the conflict, both through their effects on economic outcomes, and directly through their effects on social relations among people. Economic underdevelopment has been associated with higher conflict in the literature. Moreover, landlord areas<sup>21</sup> might still have a more unequal land distribution. “However, districts with worse land distribution historically have also seen more land reforms in the post independence period. This makes it unlikely that the persistence of the landlord effect is mainly through its effect on contemporaneous land distribution.” (Banerjee and Iyer (2005)) Also, the allocation of the responsibility of collecting land revenue

---

<sup>19</sup>Acemoglu et al. (2001, 2002), La Porta et al. (1998, 1999, 2000), Engerman and Sokoloff (1997, 2002)

<sup>20</sup>i.e. whether the country is a parliamentary democracy, or has a system of strong checks and balances.

<sup>21</sup>i.e. areas in which property rights were given to landlords. See map in FIGure B.2 in appendix B.

to the landlords (which translates into a de facto property right over the land) gave birth to a reason for perpetual conflict between the peasants and landlords. “Elsewhere, the colonial state directly collected the land revenue from the cultivator, thereby avoiding this particular source of internecine conflict.” (Banerjee and Somanathan (2007))

Moreover, land acquisition for mining purposes, building dams, or for private industry is an important issue pursued by the Maoists. Such land acquisition often leads to large scale displacement and loss of livelihood of people. Given adequate compensation people might be less discontent with the displacement induced by land acquisition. In this context, Duffo and Pande (2007) show how the landlord districts do worse than the non-landlord districts as far as effects of dams are concerned. They argue that since the social relation in the landlord districts somehow renders collective action difficult it leads to inadequate compensation. Following the same argument, if land is taken away in the landlord districts for industrial, mining or the purposes of building dams, historically people in such districts have less potential for collective action leading to inadequate compensation, which in turn leads to more grievances. This makes for easier Maoist recruits.

Thus, there are reasons to believe that the land institutions have an effect on the conflict over and above the effect through the contemporaneous land inequality and underdevelopment. We try to verify this.

**Hypothesis 5: Suitable geographic conditions impact the conflict.**

Guerrilla warfare is usually concentrated in remote rural areas where the rebels get suitable conditions to pursue such insurgencies. The favourable conditions that such regions provide include superior knowledge of local conditions and territory by rebels than by government forces, rough mountainous territory or forests providing suitable hideouts to insurgents etc. Even for nations with strong military capabilities such conditions could turn out to be daunting (e.g. the US in Vietnam, British in the initial days in Northern Ireland),<sup>22</sup> for poorer nations with weak military capabilities and corrupt bureaucratic set ups it could be hopeless.

Geography determines the cost, time and tactics necessary to curtail and check such activity. Fearon and Laitin (2003) for example find that conditions favouring insurgency like rough terrain increases the likelihood of conflict. We thus control for the geographical factors like percentage of forests, steep sloping land, sandy land, barren rocky land etc. in the district.

---

<sup>22</sup>Examples from Fearon and Laitin (2003)

Apart from the variables mentioned in the above hypotheses we control for a host of other variables that the literature has found to be important. We control for income inequality and size variables like population density, area etc.<sup>23</sup>

## 4 Empirical Analysis

Districts are the smallest administrative units in India for which coherent reliable data is available. The idea of the study is to exploit the sub-national micro structure of the data and thus all regressions are at the district level. We use data from the 16 main states in India<sup>24</sup> which covers about 90% of the Indian population. The data is pooled from diverse sources and a complete description is provided in the data appendix.

We have annual data on the conflict variables for the period 1979-2009. However, more than 90% of the data is concentrated in the years 2000 onwards and 70% in the years 2005 onwards. We divide the data into three distinct time periods due to the nature of our income data.

District-specific indicators of income or expenditure are not available for India. Instead, per capita consumption expenditure is calculated from the National Sample Survey (hereafter NSS) data. We thus use mean per capita consumption expenditure at the district level from the NSS as our proxy for income at the district level. Since we have only three rounds of NSS data viz. 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st), we club the conflict data to match these three different NSS periods. The district names are first mapped to the districts that existed in 1987. Then all the conflict data from the period 1988-1999 are collapsed and clubbed together and matched to the 1987-88 NSS data, all the conflict data from the period 2000-2004 are clubbed together and matched to the 1999-2000 NSS data, all the conflict data from the period 2005-2009 are clubbed together and matched to the 2004-2005 NSS data. Since the conflict is concentrated completely in the rural areas only rural data from the NSS is used. Thus we have data on around 360 districts for three time periods.<sup>25</sup>

There are primarily two variables of interest that we try to explain viz. ‘Probability of Conflict’ and ‘Intensity of Conflict’. In order to explain the probability of conflict we create a 0-1 binary

---

<sup>23</sup>Population might either indicate more pressure on existing resources like land or more potential recruits.

<sup>24</sup>Districts of Chattisgarh, Jharkhand and Uttaranchal are considered to be part of Madhya Pradesh, Bihar and Uttar Pradesh states respectively, i.e. their old state before the creation of the new ones.

<sup>25</sup>While running the regressions we sometimes have lesser number of districts due to the unavailability of data on certain covariates for some or all periods for certain districts.

outcome variable that takes the value ‘1’ if a district has seen any Maoist activity in the relevant period or the value ‘0’ otherwise. While it is extremely important to understand which factors increase conflict probabilities it is equally important to understand what factors lead to higher intensity of conflict. Intensity is a latent variable that is measured using two variables viz. “Total number of incidents” in a particular district and “Total number of deaths + wounded” in the district in the relevant periods. The econometric specification is straightforward:

$$(Conflict)_{j,t} = \alpha(Conflict)_{j,t-1} + \beta X_{j,t-1} + \gamma G_j + \alpha_s + \delta_t + \epsilon_{j,t}$$

The  $(Conflict)_{j,t}$  variable gives the conflict (presence or intensity) in district “ $j$ ” in round “ $t$ ”. The explanatory variables  $X_{j,t}$  include economic variables like the mean per capita consumption expenditure (MPC) and the gini coefficient of income inequality. “ $G$ ” includes all variables that do not change over time: land inequality, demographic variables, presence of marginalized factions like Scheduled Tribes and Castes and geographic variables like barren and rocky land, steep sloping land, percentage forests etc. “ $\alpha_s$ ” is the state dummy, while “ $\delta_t$ ” is the time dummy.<sup>26</sup>

For the 0-1 probability of conflict variable we use Probit specifications. The intensity variables are both count variables by nature taking integer values from zero upwards. The Poisson model is the standard model used in such cases. However, the Poisson model assumes the mean and variance to be equal. In the presence of many zeros in the data like in our case there is over-dispersion and thus the equi-dispersion assumption of the Poisson model does not hold true. The number of zero counts in our data is more than 70%. The standard parametric model to account for over-dispersion is the Negative Binomial model (Cameron and Trivedi (2005)). Thus, we use the Negative Binomial model (Poisson is a special case of it) for explaining intensity. All the regressions use cluster robust standard errors, clustered at the state level.

Our identification strategy relies on using data on the explanatory variables from the pre-conflict period. In particular, we use data on consumption expenditure (our proxy for income) from the beginning of the period. Thus, there is no serious risk of endogeneity arising out of reverse causality. Further, we always control for past conflict levels in all the specifications. However, there is still potential for endogeneity arising out of omitted variables. We thus try to control for all potential variables that might be important in our specification. However, in order to be completely sure we also run the specifications using an instrumental variables approach using

---

<sup>26</sup>The summary statistics and the correlation matrix for all the variables used in the analysis are provided in the appendix C in tables C.1 and C.2

lagged consumption expenditure as an instrument for current consumption expenditure.

## 4.1 Benchmark

Table 1 is our benchmark table. In columns 1-3, we try to understand what variables explain the presence of conflict using Probit regressions. In Columns 4-6, we run Negative Binomial regressions to explain the intensity of conflict. The dependent variable in these columns is the number of dead and wounded people in the district.

In column 1, which is our baseline specification, we see that lower incomes and higher land inequality both significantly increase the probability of conflict. A higher Scheduled Caste population does not affect the probability of conflict. However, a higher tribal population significantly increases the probability of conflict. However, this relationship is not monotonic. This is evident from the fact that the square term of the proportion Scheduled Tribe variable is negatively significant.

In column 2 we consider data from the latest 2 rounds instead of all 3 rounds and control for the initial period income using data from the first round. In column 3 we control for the state dummies in the baseline specification of column 1. Our results are robust to the different specifications.

Column 4, which is our baseline specification to explain the intensity of conflict, shows that lower incomes and higher land inequality significantly increase the number of dead and wounded people. Moreover, the presence of tribals also has a significant effect.

Like in column 2, in column 5 we consider data from the latest 2 rounds instead of all 3 and control for the initial period income using data from the first round. In column 6 we control for the state dummies in the baseline specification of column 4. The results remain unaltered.

From the above it is clear that, poorer regions experience more conflict. This is in line with the previous literature and is in fact one of the most robust results in the conflict literature. In terms of average marginal effects, a 1 standard deviation increase in the log of mean per capita consumption expenditure increases the probability of conflict in the next period by 5%, and results in 50 more dead and wounded people in a year.

On the other hand, while income inequality does not seem to play any role in the context of the Maoist conflict, land inequality is a highly significant and robust predictor of conflict in all

Table 1: The determinants of the Maoist Conflict in India

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
consumption per capita	-1.323*** (0.425)	-1.434*** (0.528)	-1.561** (0.734)	-3.014** (1.478)	-1.530 (1.272)	-4.239** (1.922)
land inequality	3.617*** (0.781)	3.566*** (1.202)	4.435*** (0.899)	15.01*** (2.934)	13.52*** (3.083)	14.68*** (3.647)
proportion sandy	-1.524 (2.645)	-0.603 (2.687)	0.332 (1.393)	-31.58 (21.01)	-42.51** (21.55)	-10.25 (8.838)
log state capital distance	-0.0823 (0.147)	0.0521 (0.146)	-0.0201 (0.178)	-0.795*** (0.216)	-0.472** (0.190)	-0.303 (0.219)
proportion barrenrocky	10.50*** (2.734)	15.93*** (3.510)	13.12*** (4.414)	40.97** (16.73)	47.68*** (16.70)	26.93*** (4.656)
proportion steepslping	-67.35*** (21.16)	-86.49*** (32.92)	-126.6*** (32.94)	-270.3*** (36.59)	-242.1*** (56.58)	-346.3*** (72.58)
proportion forest cover	0.889 (0.708)	1.568** (0.775)	2.397*** (0.597)	2.669 (2.040)	2.965 (2.075)	6.060*** (1.757)
%Scheduled Castes	-0.629 (1.349)	0.0471 (1.517)	2.922** (1.291)	-4.485 (4.095)	-5.038* (2.794)	7.018* (4.046)
%Scheduled Tribes	4.383*** (1.651)	4.229** (1.872)	2.882*** (0.677)	10.04** (4.432)	10.01*** (3.621)	6.231*** (2.243)
%Scheduled Tribes square	-5.630** (2.805)	-4.826** (2.309)	-2.062 (1.361)	-12.15** (6.174)	-13.49*** (4.954)	-3.831 (2.832)
population density	0.293 (0.230)	0.618* (0.329)	0.329 (0.297)	-0.150 (0.562)	-0.00329 (0.515)	-0.0382 (0.535)
log area	0.671*** (0.216)	0.749*** (0.275)	0.681*** (0.142)	2.416*** (0.408)	2.380*** (0.394)	2.388*** (0.396)
income inequality	-0.131 (2.057)	-3.545* (1.849)	1.658 (1.500)	-7.344 (5.025)	-14.11** (5.544)	-1.936 (3.095)
initial consumption per capita		-0.491 (0.615)			-3.749*** (0.778)	
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	996	656	996	996	656	996
ll	-269.1	-181.9	-192.9	-1013.1	-829.0	-913.4

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

the specifications. Land inequality also significantly increases the intensity of conflict and this is robust to different controls. This gives support to the grievance arising out of land inequities being a key reason behind the conflict. In terms of average marginal effects if we look at our baseline specification, a 1 standard deviation increase in land inequality increases the probability of conflict in the next period by 9%, and results in 157 more dead and wounded people a year.

As mentioned above the percentage of Scheduled Tribes in the district has a non monotonic

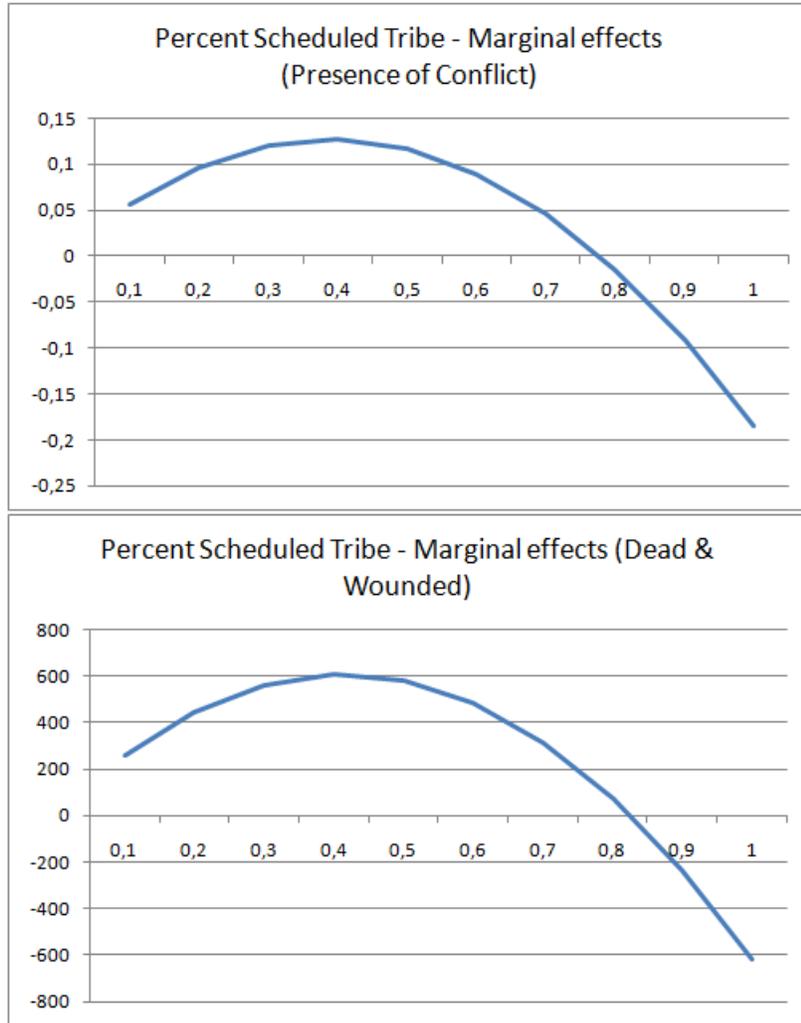


Figure 1: Effects of Scheduled Tribe presence

impact on conflict. In Figure 1, we notice that the marginal effect increases till about 40% tribal presence in the district and then it starts falling. At around 80% tribal presence the marginal effect actually becomes negative. We see that districts which have around 40% more Scheduled Tribes face around 12% higher probability of conflict and around 120 more dead and wounded people in a year.

As far as the geography variables are concerned, they turn out to be quite important. We see that barren and rocky areas experience more conflict. This is in agreement with our expectations. In a primarily agrarian economy, a higher proportion of barren and rocky land indicates

lack of economic opportunity. This supports the opportunity cost story.

The percentage of forest cover variable comes out to be significantly increasing the probability and intensity of conflict in many of the specifications (but not all). One explanation could be that large patches of forest cover provides the perfect hiding and fighting conditions for rebels and makes it quite difficult for security forces to keep up with them. Also, we see that districts with higher geographical areas experience both a higher probability and intensity of conflict.

The percentage steep sloping land variable on the other hand comes out to be negatively significant. This could sound a bit counter intuitive since most of the previous literature mentions mountainous land as being more suitable for conflict. However, since the conflict is concentrated primarily in the plains of India the percentage of mountainous land in a district might actually have a negative sign.

In Table 2 we add a control for historical institutions to the otherwise identical specifications of Table 1. The relevant variable gives the proportion of the district that was not under the control of the landlords. The historical institutions variable has a significant and quite robust impact on the conflict. However, consumption per capita becomes less robust while land inequality on the other hand continues to be significant and robust. The presence of Scheduled Tribes does not seem to matter anymore. However, by adding the control for historical institutions we lose a third of the observations due to the unavailability of data on the institutions variable for all districts.<sup>27</sup> Thus, sample change might be the cause of the change in significance of the other variables. In fact, this is what seems to be the reason behind the Scheduled Tribe and the consumption per capita variables turning insignificant in Table 2.<sup>28</sup>

In Table 3, we instrument current consumption expenditure with its own lagged values. Otherwise, we have the identical specifications of Table 1. We use two stage instrumental variable regressions. In the first stage we notice that the instrument is highly significant.

In the second stage we notice a few changes from Table 1. We find that consumption expenditure per capita and land inequality continue to be significant in explaining the probability and the intensity of conflict. However, we notice that the presence of Scheduled Tribes is no longer significant. We also report the coefficients of the proportion Scheduled Tribe variable in the

---

<sup>27</sup>Data on the land institutions variable is available only for the districts which were directly under British control.

<sup>28</sup>See Table C.3 in appendix C where we have the same sample as in Table 2.

first stage. We see that the proportion Scheduled Tribe significantly reduces the per capita consumption expenditure. Thus, in the second stage once we control for this effect, the proportion Scheduled Tribe variable does not matter anymore.

Table 2: The role of historical institutions

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
consumption per capita	-0.862* (0.470)	-0.955* (0.556)	-0.942 (0.891)	-1.724 (1.294)	-1.378 (1.131)	-3.115 (2.809)
land inequality	2.438*** (0.546)	2.170** (0.852)	3.917*** (0.901)	11.20*** (2.189)	9.867*** (2.359)	13.36*** (2.998)
proportion Non landlord	-0.935*** (0.276)	-1.041*** (0.249)	-1.188** (0.476)	-2.496*** (0.679)	-2.088*** (0.609)	-1.402 (0.859)
proportion sandy	17.66** (7.832)	20.83*** (6.724)	32.15** (15.35)	44.76 (78.21)	-17.15 (38.28)	42.28 (55.40)
log state capital distance	0.190 (0.195)	0.243 (0.219)	0.163 (0.208)	-0.0216 (0.470)	0.134 (0.352)	-0.127 (0.393)
proportion barrenrocky	19.02*** (3.395)	26.28*** (4.609)	22.95*** (6.081)	54.94** (25.72)	51.73*** (15.01)	39.80*** (9.432)
proportion steepsloping	-53.55* (27.34)	-95.53*** (35.33)	-132.3*** (43.57)	-207.3*** (74.18)	-331.6*** (116.1)	-366.8*** (101.9)
proportion forest cover	1.543** (0.692)	2.109*** (0.695)	2.727*** (0.622)	5.319*** (1.922)	5.458*** (1.842)	7.333*** (1.904)
%Scheduled Castes	-0.969 (1.398)	-0.592 (1.520)	4.252*** (1.310)	-4.802 (4.316)	-6.595 (4.504)	10.62*** (3.125)
%Scheduled Tribes	2.749 (2.542)	4.273* (2.596)	1.438 (1.160)	-3.397 (6.369)	-1.640 (7.405)	-0.123 (3.261)
%Scheduled Tribes square	-5.565 (4.337)	-8.690** (4.050)	-1.199 (2.020)	-2.575 (10.20)	-8.935 (12.91)	2.205 (4.054)
population density	0.484** (0.203)	0.762** (0.341)	0.413 (0.317)	0.199 (0.597)	0.0307 (0.651)	0.134 (0.642)
log area	0.846*** (0.200)	1.008*** (0.317)	0.902*** (0.138)	3.316*** (0.587)	3.029*** (0.498)	2.782*** (0.326)
income inequality	-2.807 (1.865)	-5.845*** (1.724)	-1.116 (1.937)	-15.47*** (5.244)	-20.54*** (6.346)	-3.583 (6.438)
initial consumption per capita		0.242 (0.811)			-2.315*** (0.765)	
State dummies			Yes			Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	655	431	655	655	431	655
ll	-177.9	-131.6	-138.2	-642.2	-535.5	-590.7

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

This table is identical to Table 1 but also controls for historical institutions. The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Table 3: The determinants of the Maoist Conflict in India - IV estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
consumption per capita	-2.642*** (0.857)	-3.033*** (1.164)	-4.556*** (1.160)	-6.971*** (1.651)	-4.813** (1.988)	-10.34*** (3.006)
land inequality	3.171*** (1.103)	3.183** (1.367)	3.768*** (0.992)	12.78*** (3.440)	13.48*** (3.620)	7.892*** (2.894)
proportion sandy	0.516 (2.530)	0.443 (2.926)	0.397 (1.567)	-25.68* (15.35)	-35.76** (16.76)	-12.84 (11.73)
log state capital distance	0.0129 (0.153)	0.0316 (0.141)	0.0567 (0.158)	-0.616*** (0.180)	-0.546*** (0.177)	-0.123 (0.220)
proportion barrenrocky	15.11*** (3.832)	16.40*** (3.592)	14.98** (6.119)	49.91*** (14.19)	52.07*** (14.49)	32.64*** (8.903)
proportion steepslowing	-48.50** (22.01)	-81.30** (33.72)	-91.21** (39.60)	-174.1*** (25.66)	-232.2*** (62.55)	-232.2*** (62.66)
proportion forest cover	1.642* (0.855)	1.597** (0.804)	2.872*** (0.854)	3.855* (2.217)	3.562 (2.180)	7.583*** (1.974)
%Scheduled Castes	-0.249 (1.635)	-0.430 (1.677)	2.022* (1.084)	-6.754** (2.914)	-6.538** (2.847)	0.573 (3.741)
%Scheduled Tribes	2.809 (2.196)	2.486 (2.293)	-0.766 (2.136)	2.963 (2.874)	5.844 (3.693)	-1.733 (2.291)
%Scheduled Tribes sq	-3.444 (2.819)	-3.043 (2.816)	1.693 (2.703)	-5.680 (3.765)	-9.201** (4.620)	1.831 (2.841)
population density	0.577 (0.356)	0.535 (0.371)	0.258 (0.382)	-0.357 (0.628)	-0.303 (0.639)	0.315 (0.711)
log area	0.628** (0.311)	0.608** (0.297)	0.464** (0.207)	1.793*** (0.358)	2.112*** (0.343)	1.816*** (0.415)
income inequality	-1.296 (2.200)	-0.516 (2.595)	5.412** (2.418)	-6.500 (5.021)	-10.22* (5.981)	7.342 (5.337)
initial consumption per capita	No	Yes	No	No	Yes	No
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
First Stage						
consumption per capita lagged	<b>0.602***</b> (0.0642)	<b>0.461***</b> (0.0608)	<b>0.359***</b> (0.0520)	<b>0.601***</b> (0.065)	<b>0.461***</b> (0.065)	<b>0.358***</b> (0.052)
%Scheduled Tribes	-0.679*** (0.166)	-0.674*** (0.170)	-0.684*** (0.136)	-0.679*** (0.168)	-0.678*** (0.171)	-0.680*** (0.140)
%Scheduled Tribes square	0.636*** (0.158)	0.656*** (0.165)	0.678*** (0.174)	0.636*** (0.166)	0.665*** (0.170)	0.669*** (0.180)
Observations	640	638	640	640	638	640
ll	84.31	91.60	236.0	-834.1	-826.5	-762.2

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using IV-Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using OLS regressions in the first stage and Negative Binomial regressions in the second stage. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds. Previous period consumption expenditure instruments following period consumption expenditure.

## 4.2 Testing for the “Exclusion” hypothesis

In Table 4, we check specifically for the exclusion story. We run the same set of regressions as in Table 1 but instead of the overall per capita income of the district as the explanatory variable we use the per capita incomes of the General Castes, Scheduled Castes and Scheduled Tribes as three different explanatory variables within the same specification. This allows us to verify if the income of a specific ethnic group is important for the conflict while controlling for the incomes of the other groups. One limitation with this analysis is that the number of Scheduled Tribal people surveyed is low in many of the districts owing to the low percentage of Scheduled Tribes living in those districts. We include only districts which have at least a weighted population of 15 tribal people surveyed.<sup>29</sup> This significantly reduces our sample size. However, the comparison of incomes across ethnic groups is suitable only in districts which have at least some residents of each of the groups we would like to compare, actually residing in those districts. Thus, we proceed with our analysis with the reduced sample size. We do not find evidence of conflict arising out of lower incomes of the Scheduled Tribes vis-a-vis incomes of the General Castes and others. In fact, it is the income of the General Castes that has a significant negative effect on the conflict. Also, in three of the six specifications the incomes of the Scheduled Castes come out to be significantly reducing the conflict. But we must be cautious in interpreting this result. This only implies that as far as levels of income are concerned, we do not find evidence on tribal incomes being important for the conflict. However, (as we will see later) once we analyze differences in growth rates of income across ethnic groups we might get a completely different story.<sup>30</sup>

The lower castes and the tribes are ethnically two very different groups of people. However, both these groups are seriously disadvantaged socially and economically. In order to be able to use more observations, we club the Scheduled Castes and Tribes together and treat them as one group <sup>31</sup>. We find that the incomes of both groups matter. However, the coefficient for the incomes of the Scheduled Castes and Tribes is almost always greater than that of General Castes. Moreover, in almost all the specifications we have percentage of Scheduled Castes and Tribes in the district, significantly increasing the conflict presence and intensity. While this is

---

<sup>29</sup>Including districts which have even fewer or more Scheduled Tribal people surveyed makes no significant difference in the results.

<sup>30</sup>In Table C.5 we control for the historical institutions variable alongside incomes of the different ethnic groups. There is no qualitative change in the results.

<sup>31</sup>See Table C.4

not conclusive, it does give some confirmation of the participation from the lower castes and tribes.

Table 4: The role of income differences across ethnic groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
General Castes consumption pc	-0.818** (0.319)	-0.990*** (0.363)	-1.273*** (0.202)	-1.871** (0.874)	-1.951*** (0.756)	-2.109*** (0.680)
Scheduled Castes consumption pc	-0.444 (0.348)	-0.743*** (0.245)	-0.675*** (0.186)	-1.324 (0.832)	-1.338 (0.831)	-1.874* (0.986)
Scheduled Tribes consumption pc	-0.218 (0.484)	-0.135 (0.578)	-0.117 (0.695)	-0.670 (1.407)	-0.814 (1.223)	-0.690 (1.051)
proportion sandy	-3.882 (5.447)	-1.202 (5.517)	3.041 (3.221)	-83.26*** (30.86)	-84.37*** (29.02)	-34.19*** (9.921)
log state capital distance	-0.141 (0.163)	-0.00183 (0.118)	-0.0944 (0.250)	-0.916*** (0.247)	-0.695** (0.323)	-0.458** (0.194)
proportion barrenrocky	7.414* (4.261)	16.08*** (5.217)	15.35*** (4.608)	29.98 (21.45)	28.05 (23.02)	17.40*** (4.523)
proportion steepsloping	-115.0*** (39.00)	-116.2** (55.10)	-156.6** (60.87)	-347.1*** (126.2)	-296.4*** (101.8)	-338.8*** (131.5)
proportion forest cover	0.846 (0.891)	2.221** (1.036)	3.682*** (0.728)	1.011 (1.287)	0.141 (1.691)	6.027*** (1.437)
%Scheduled Castes	-0.0988 (1.420)	0.504 (1.185)	2.450 (1.839)	-3.464 (4.206)	-6.589* (3.377)	4.253 (5.513)
%Scheduled Tribes	2.080 (1.369)	1.098 (2.007)	1.576 (1.472)	4.489 (3.617)	2.564 (4.295)	0.210 (2.729)
%Scheduled Tribes square	-1.991 (2.278)	-0.00427 (2.286)	0.158 (2.294)	-3.798 (3.626)	-1.908 (4.174)	2.840 (3.320)
population density	-0.166 (0.379)	0.344 (0.363)	0.351 (0.368)	-1.684 (1.040)	-1.549 (1.092)	-1.117 (0.737)
log area	0.806*** (0.198)	0.949*** (0.282)	0.828*** (0.208)	2.784*** (0.379)	2.553*** (0.383)	2.536*** (0.315)
income inequality	1.504 (1.635)	-0.945 (2.433)	2.555* (1.552)	-5.349 (3.958)	-7.974* (4.254)	-5.269 (4.035)
land inequality	5.412*** (0.902)	5.575*** (1.062)	4.555*** (0.841)	20.47*** (5.315)	18.98*** (5.471)	16.20*** (4.213)
initial consumption per capita		-0.788* (0.423)			0.163 (1.013)	
State dummies			Yes			Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	567	356	552	567	356	567
ll	-163.0	-99.64	-118.5	-761.9	-616.9	-698.6

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

### 4.3 Growth and Conflict

Table 5: The Effect of Growth on Conflict

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
Consumption growth	-0.0213 (0.0181)	-0.0335* (0.0183)	-0.00481 (0.0181)	0.0365 (0.0371)	-0.00927 (0.0389)	-0.0203 (0.0408)
proportion sandy	-0.991 (2.787)	-1.386 (2.807)	0.488 (1.521)	-31.55* (17.17)	-42.08** (20.78)	-16.08 (13.19)
log state capital distance	0.0572 (0.161)	0.0510 (0.150)	0.167 (0.195)	-0.527** (0.207)	-0.492** (0.191)	0.102 (0.272)
proportion barrenrocky	12.03** (5.028)	15.55*** (3.658)	13.29** (6.089)	47.66*** (15.12)	52.54*** (14.98)	27.48*** (9.349)
proportion steepsloning	-69.79*** (22.75)	-92.50*** (31.15)	-97.17*** (37.30)	-194.3*** (30.34)	-247.5*** (68.15)	-248.7*** (70.50)
proportion forest cover	1.094 (0.812)	1.526** (0.732)	2.430*** (0.782)	2.730 (2.373)	3.006 (2.187)	5.793*** (1.982)
%Scheduled Castes	0.296 (1.201)	0.219 (1.399)	3.189** (1.479)	-3.939 (3.478)	-4.777* (2.841)	4.423 (3.665)
%Scheduled Tribes	5.777*** (1.550)	4.969*** (1.786)	3.853*** (0.835)	10.47*** (3.743)	10.78*** (3.135)	9.167*** (3.442)
%Scheduled Tribes square	-5.987*** (1.971)	-5.526** (2.170)	-2.844* (1.726)	-10.77** (4.668)	-13.80*** (4.550)	-8.308 (5.261)
population density	0.727*** (0.258)	0.619** (0.304)	0.649* (0.373)	0.108 (0.689)	-0.0894 (0.643)	0.858 (0.792)
log area	0.871*** (0.278)	0.784*** (0.270)	0.913*** (0.199)	2.367*** (0.323)	2.432*** (0.329)	2.687*** (0.575)
income inequality	-6.188*** (1.951)	-4.961*** (1.883)	-2.238** (1.058)	-21.51*** (4.989)	-17.64*** (5.166)	-10.17*** (3.142)
land inequality	3.561*** (0.970)	3.686*** (1.164)	4.187*** (1.028)	13.81*** (2.946)	14.21*** (3.278)	9.756*** (3.178)
initial consumption per capita		-1.401** (0.608)			-4.263*** (0.932)	
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	640	638	640	640	638	640
ll	-193.5	-185.3	-143.5	-844.2	-828.3	-770.9

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The growth rates correspond to the 2 latter rounds. The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

In Table 5, we try to identify what impact growth might have on the conflict. Instead of the district income level we now include the district income growth rate as the explanatory variable. We see that growth does not have any significant impact on the conflict (We see this in five out of the six specifications). Moreover, the average marginal effects are also small in magnitude. As far as the other variables are concerned, land inequality and the presence of

Table 6: The growth in income of different ethnic groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Intensity	Intensity	Intensity	Intensity
General Castes growth	-0.0263** (0.0129)	-0.0325** (0.0149)	0.00549 (0.0168)	-0.00239 (0.0160)	0.00901 (0.00972)	-0.00426 (0.00835)
Scheduled Castes growth	-0.0142 (0.0137)	0.00539 (0.0105)	0.0377 (0.0247)	0.0386* (0.0227)	0.0142 (0.0121)	0.0244 (0.0155)
Scheduled Tribes growth	-0.0156 (0.0187)	-0.0201 (0.0276)	<b>-0.0930***</b> (0.0290)	<b>-0.0695*</b> (0.0355)	<b>-0.0595***</b> (0.0231)	<b>-0.0438**</b> (0.0220)
proportion sandy	-4.741 (5.887)	-0.0125 (2.860)	-84.78*** (32.08)	-39.39** (17.89)	-23.26** (10.71)	-13.07*** (4.514)
log state capital distance	0.118 (0.104)	0.129 (0.211)	-0.347 (0.220)	0.0230 (0.323)	-0.303 (0.227)	0.00305 (0.263)
proportion barrenrocky	15.97*** (5.749)	21.55*** (8.341)	30.53 (26.10)	27.05*** (9.680)	21.27 (14.81)	21.06** (9.082)
proportion steepslping	-185.2*** (53.76)	-251.8*** (95.98)	-442.7*** (129.8)	-352.4*** (130.7)	-241.6*** (72.69)	-198.3*** (74.77)
proportion forest cover	2.461*** (0.788)	4.882*** (1.418)	-1.162 (1.552)	4.502 (2.788)	1.365 (1.539)	5.412*** (1.923)
%Scheduled Castes	0.950 (1.353)	3.568 (2.990)	-7.556* (3.952)	1.666 (5.108)	-1.812 (3.485)	3.120 (4.505)
%Scheduled Tribes	1.548 (2.319)	0.638 (1.816)	8.996 (5.636)	6.967*** (2.675)	4.568 (3.725)	2.847 (1.978)
%Scheduled Tribes square	-1.006 (2.683)	0.231 (2.076)	-12.87* (6.687)	-8.309*** (2.856)	-5.967 (4.325)	-1.821 (2.252)
population density	0.192 (0.438)	0.264 (0.490)	-0.701 (0.749)	0.0926 (0.921)	-0.0870 (0.628)	0.663 (0.710)
log area	1.242*** (0.319)	0.996*** (0.255)	3.705*** (0.478)	3.152*** (0.542)	2.657*** (0.619)	2.370*** (0.500)
income inequality	-5.516** (2.189)	-6.117*** (1.658)	-15.73*** (4.707)	-14.37*** (5.475)	-11.84*** (3.053)	-9.873** (4.345)
land inequality	7.679*** (1.436)	7.624*** (0.894)	19.09*** (3.899)	12.58*** (4.120)	12.09*** (2.270)	7.446*** (2.569)
initial consumption per capita	No	Yes	No	Yes	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
State dummies	No	Yes	No	Yes	No	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	290	284	290	290	290	290
ll	-74.40	-57.41	-544.7	-511.9	-493.8	-453.0

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1 and 2 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 3, 4, 5 and 6 explain the intensity of conflict (columns 3 and 4 - the no. of dead & wounded in the district; columns 5 and 6 - the number of incidents in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The growth rates correspond to the 2 latter rounds. The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Scheduled Tribes are significant and robust in all the specifications. <sup>32</sup>

<sup>32</sup>In Table C.6 we control for the historical institutions variable alongside growth. There is no qualitative change in the results.

In Table 6, we try to identify if horizontal inequality in growth i.e. differences in the growth of incomes of the different ethnic groups have any significant impact on the conflict. The dependent variable in columns 1 and 2 is the presence of conflict. In columns 3 to 6 it is the intensity of conflict.<sup>33</sup> We see that as far as the presence of conflict is concerned (i.e. columns 1 and 2), it is the growth rate of the incomes of the General Castes that matter. In terms of marginal effects, a 1 standard deviation increase in the growth rate for the General Castes leads to a 4% lower probability of conflict while a similar increase for the Scheduled Castes and Tribes is not significant in the baseline specification of column 1. However, for the specifications explaining intensity, the growth rates of the General Castes is not significant and even changes its sign in some of the specifications. On the other hand a growth in incomes of the Scheduled Tribes significantly and robustly reduces the intensity of conflict in all the specifications (columns 3-6). The growth rate for Scheduled Castes does not seem to be important. In terms of marginal effects, a 1 standard deviation fall in the income growth of the Scheduled Tribes leads to an increase in the number of dead and wounded people in the district by around 155 people in the next period. Moreover, this relation is highly significantly and robust.

Thus, from the results in Table 6 we do get evidence for the “Exclusion hypothesis”. For the presence of conflict the growth in incomes of the General Castes is significant, but the marginal effects are small and are similar in magnitude to the effects of growth rates of other ethnic groups. However, the conflict is significantly more intense in districts which face lower growth in the incomes of the Scheduled Tribes and these effects are robust. This means that some of the participation in the conflict by the tribals and/or support for the Maoists among them might actually be driven by the lower growth in incomes of the Scheduled Tribes as they feel excluded from the growth that the rest of society is facing.

#### 4.4 Temporal variation in the conflict

In the previous tables we have found some interesting insights about the causes of the Naxalite conflict in India. However, it will be interesting to check whether there is a change in the nature and causes of the conflict over time. In Table 7, we thus regress the conflict variables over the 3 rounds separately. This table gives us some more interesting insights. Columns 1-3 show

---

<sup>33</sup>In columns 3 and 4 the dependent variable is the number of dead and wounded people and in columns 5 and 6, it is the number of incidents in the district.

the change in causes of the presence of conflict over the three rounds, while columns 4-6 show the corresponding change in the causes of intensity of conflict. We see that land inequality is important across all periods. But its relative importance varies over time. We see that to start with land inequality is a very important factor but over time the magnitude of the coefficient falls, suggesting that its relative importance is declining over time.

On the other hand we see that to start with Consumption expenditure is not that important, but in the first half of decade of 2000s (round 2) the conflict shifts to poorer areas. And finally we see that the conflict actually in the last five years (round 3) moves to areas where there are more tribal people. However, as in Table 1, this relationship is non-monotonic. We also see how in the first round the proportion of barren and rocky land does not matter but it does in the next rounds.

Interestingly we see that initially having a higher proportion of tribals in the district does not matter. In rounds 1 and 2 the proportion of Scheduled Tribes in the district is not significant. However, in round 3 we notice that having a higher proportion of tribals in the district significantly increases the conflict. This perhaps goes to show that in recent times the conflict has been spreading to tribal areas.

We also try to split the incomes of the separate groups and try to identify the temporal differences in their effects. We do not find any significant results. This is perhaps driven by the fact that we lose a lot of observations in the rounds 2 and 3 which are in fact the rounds where most of the conflict is concentrated.<sup>34</sup>

---

<sup>34</sup>See Table C.7

Table 7: Change in causes of conflict over time

	Round 1 Probability	Round 2 Probability	Round 3 Probability	Round 1 Intensity	Round 2 Intensity	Round 3 Intensity
consumption per capita	0.698 (0.663)	-3.436*** (0.590)	-0.990 (0.673)	0.875 (3.301)	-6.042*** (1.602)	-2.333 (1.509)
land inequality	4.796*** (1.290)	3.726*** (1.221)	3.605** (1.447)	20.19*** (2.987)	18.29*** (4.070)	10.32*** (2.835)
%Scheduled Tribes	3.219 (3.656)	-0.652 (2.438)	8.317*** (2.126)	10.16 (11.80)	7.630* (4.375)	10.66** (5.202)
%Scheduled Tribes square	-5.395 (6.475)	0.930 (3.454)	-10.31*** (2.776)	-16.05 (14.10)	-7.126 (5.571)	-17.65** (7.297)
%Scheduled Castes	-2.400 (2.472)	-3.382** (1.601)	1.916 (1.923)	-5.941 (7.436)	-4.721 (3.660)	-4.695 (3.622)
proportion sandy	-27.77 (19.41)	0.125 (2.847)	0.547 (2.561)	-98.65* (59.65)	-150.7*** (38.47)	-12.81 (17.52)
proportion barrenrocky	-0.203 (4.344)	22.24*** (4.492)	10.42* (5.641)	56.20 (44.11)	32.23** (14.71)	41.43** (16.58)
proportion steepslowing	-139.3*** (48.60)	-134.4*** (40.72)	-36.40* (19.22)	-829.2*** (230.3)	-202.2*** (44.76)	-173.5*** (30.76)
log state capital distance	-0.330* (0.176)	0.178 (0.157)	-0.0338 (0.257)	-0.748* (0.384)	-0.690*** (0.261)	-0.376* (0.222)
proportion forest cover	-0.172 (1.315)	1.866* (1.087)	1.368 (1.125)	3.163 (2.656)	-0.458 (1.138)	3.479 (2.500)
log area	1.208*** (0.335)	0.927*** (0.356)	0.574** (0.276)	6.459*** (1.085)	2.569*** (0.480)	1.773*** (0.426)
population density	-0.316 (0.474)	0.679* (0.391)	0.623 (0.421)	1.239 (1.168)	-1.219 (0.831)	0.542 (0.546)
income inequality	5.712* (3.365)	-1.404 (1.258)	-4.271** (2.118)	0.758 (10.35)	-15.05 (9.751)	-9.497* (5.252)
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	336	321	339	336	321	339
ll	-61.32	-85.11	-88.77	-147.0	-306.9	-522.6

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

## 5 Conclusion

This paper studies the political economy of the Maoist conflict in India in-depth using a district level panel. It also contributes to the civil conflict literature by adding to a growing number of studies that use sub-national micro data to study civil conflicts. Making use of a newly constructed district level conflict database the paper provides some interesting insights on the causes of the Maoist conflict in India.

The main finding of the paper is that the story behind the Maoist Conflict in India is a story of grievances arising out of feelings of exclusion of various forms. We see how the underdeveloped districts and districts with higher land inequality are more prone to conflict. Moreover, we have evidence in favour of participation from the Scheduled Tribes in the conflict. We see how low growth rates in the incomes of Scheduled Tribes leads to more conflict. Also, we see how social divisions created by historical institutions play a role in the conflict. All these findings indicate that certain sections of society feel left behind in the growth story of India leading to more grievances and social tensions which help Maoists by creating more sympathizers and boosting their recruitment efforts.

Future research would be in the direction of controlling for land reforms at the district level and verifying its effects on the conflict outcomes. The other direction for future research would be to find adequate instruments for the income variables that would allow us to further ensure that our results are not affected by endogeneity. Moreover, a lot of work remains to be done in terms of data collection from the households of the perpetrators and victims in order to further pin down both the causes and consequences of the Maoist conflict at the household/individual level.

## References

- Daron Acemoglu, Simon Johnson, and James A. Robinson. The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91 (5):1369–1401, 2001.
- Daron Acemoglu, Simon Johnson, and James A. Robinson. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics*, 117 (4):1231–94, 2002.
- Catherine Andre and Jean-Philippe Platteau. Land tenure under unendurable stress: Rwanda caught in the malthusian trap. *Cahiers de la Faculte des Sciences Econommiques et Sociales de Namur, Serie Recherche*, 164 - 1996/7, 1996.
- Abhijit Banerjee and Laxmi Iyer. History, institutions and economic performance: The legacy of colonial land tenure systmes in india. *American Economic Review*, 2005.
- Abhijit Banerjee and Rohini Somanathan. The political economy of public goods: Some evidence from india. *Journal of Development Economics, Elsevier*, 82(2):287–314, 2007.
- Vani K. Barooah. Deprivation, violence and conflict: An analysis of naxalite activity in the districts of india. *International Journal of Conflict and Violence*, 2 (2):317–333, 2008.
- Gary S Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217, 1968.
- Timothy Besley and Robin Burgess. Land reform, poverty reduction, and growth: Evidence from india. *Quaterly Journal of Economics*, 115 (2):389–430, 2000.
- Timothy J. Besley and Torsten Persson. The incidence of civil war: Theory and evidence. *NBER Working Paper Series, Working Paper 14585*, 2008.
- Bela Bhatia. The naxalite movement in central bihar. *Economic and Political Weekly*, 40 (15) Apr 9-15, 2005.
- C. Blattman and E. Miguel. Civil war. *The Journal of Economic Literature*, 48(1):3–57(55), 2010.

- A. Colin Cameron and Pravin K. Trivedi. *Microeconometrics:Methods and Applications*. 2005.
- Sudeep Chakravarti. Red sun-travels in naxalite country. *Penguin Global*, 2008.
- Antonio Ciccone. Transitory economic shocks and civil conflict. *unpublished*, 82(2):287–314, 2010.
- P. Collier and A. Hoeffler. Greed and grievance in civil war. *Oxford Economic Papers*, 56(4): 563–596, 2004.
- Quy-Toan Do and Lakshmi Iyer. Geography, poverty and conflict in nepal. *HBS Working Paper*, 07-065, 2009.
- Oeindrila Dube and Juan F. Vargas. Commodity price shocks and civil conflict:evidence from colombia. *CID Graduate Student and Postdoctoral Fellow Working Paper No. 14*, 2008.
- Esther Duflo and Rohini Pande. Dams. *The Quarterly Journal of Economics*, 122(2):601–646, 05, 2007.
- Stanley L. Engerman and Kenneth L. Sokoloff. Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the united states. in *Steven Haber, ed., How Latin America fell behind: Essays on the economic histories of Brazil and Mexico, 18001914*. Stanford: Stanford University Press, 1997.
- Stanley L. Engerman and Kenneth L. Sokoloff. Factor endowments, inequality, and paths of development among new world economies. *Economia: Journal of the Latin American and Caribbean Economic Association*, 3 (1):41–88, 2002.
- J. Fearon and D. Laitin. Ethnicity, insurgency and civil war. *American Political Science Review*, 97 (1):75–90, 2003.
- Kishore Gawande, Devesh Kapur, and Shanker Satyanath. Natural resource shock and conflict in india’s red belt. *Work in progress*, 2012.
- Ramachandra Guha. Adivasis,naxalites and indian democracy. *Economic and Political Weekly (Special Articles)*, August 11, 2007.
- John Harris. The naxalite/maoist movement in india: A review of recent literature. *ISAS Working paper*, No. 109, 8 July, 2010.

- Kristian Hoelscher, Jason Mikilian, and Krishna Chaitanya Vadlamannati. Hearts and mines: A district-level analysis of the maoist conflict in india. *available at www.uni-heidelberg.de/md/awi/professuren/intwipol/india.pdf*, 2011.
- Lakshmi Iyer. The bloody millennium: Internal conflict in south asia. *Harvard Business School, Working Paper*, 09-086, 2009.
- Saumitra Jha. Complementarities and religious tolerance: evidence from india. *Stanford GSB Working Paper No. 2004*, 2008.
- Rajat Kujur. Naxal movement in india: A profile. *Institute of Peace and Conflict Studies, New Delhi*, 2008.
- Rafael La Porta, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny. Law and finance. *Journal of Political Economy*, 106 (6):113–55, 1998.
- Rafael La Porta, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny. The quality of government. *Journal of Law, Economics, and Organization*, 15 (1):222–79, 1999.
- Rafael La Porta, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny. Investor protection and corporate governance. *Journal of Financial Economics*, 58 (1-2):3–27, 2000.
- Karen Macours. Increasing inequality and civil conflict in nepal. *Oxford Economic Papers*, 63(1): 1–26, 2011.
- E. Miguel, S. Satyanath, and E. Sergenti. Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112 (41):725–753, 2004.
- Anirban Mitra and Debraj Ray. Implications of an economic theory of conflict:hindu-muslim violence in india. *unpublished*, 2010.
- Government of India. Development challenges in extremist affected areas. *Report of an Expert Group to the Planning Commission*, 2008.
- Rohini Pande. Can mandated political representation increase policy influence for disadvantaged minorities? theory and evidence from india. *American Economic Review*, 93, No.4:1132–1151, 2003.
- Nandini Sundar. Subalterns and sovereigns, an anthropological history of bastar (1854-2006). *Oxford University Press*, 2008.

Oliver Vanden Eynde. Targets of violence: Evidence from india's naxalite conflict. *Job market paper, LSE*, 2011.

Philip Verwimp. Micro-level evidence from rwanda. *HiCN Working Papers 08, Households in Conflict Network*, 2003.

## A Data appendix

### Naxalite/Maoist incidents:

The data on the Maoist incidents comes from four different sources:

- Global Terrorism Database (GTD) I: 1970-1997 & II: 1998-2004: In fact now there is one consolidated GTD which has data till 2007.
- Rand-MIPT Terrorism Incident database (1998- present)
- Worldwide Incidents tracking system (WITS), National Counter Terrorism Centre (2004-2007)
- South Asia Terrorism Portal (SATP) maintains a comprehensive web portal tracking all terrorism incidents in South Asia. It has most detailed accounts of the major Naxalite incidents from 2005 onwards.

All the above mentioned sources track data on terrorist/violent incidents from different sources like newspapers, official reports etc.<sup>35</sup> The data is across countries often with the village/city/district name of the place where each particular incident took place specified. We have filtered out all the data pertaining to India and then kept only the incidents that were clearly identifiable as Maoist/Naxalite in nature. To do so we used the name of perpetrator unless it was clearly mentioned as being a Maoist/Naxalite incident. Then except for the cases where the district is clearly mentioned the town or village mentioned was placed in its corresponding district. More often than not we were successful in doing so. However, we had to leave out the cases where it was not possible to do identify the districts. Finally, we have consolidated all the above data sources to construct a comprehensive consolidated district level Maoist incidents database. Thus, there is information on presence, incidents and the number of deaths and injuries at the district level from 1979-2009.

**Consumption:** District-specific indicators of income or expenditure are not available for India. Instead, per capita consumption expenditure is calculated from the National Sample Survey (hereafter NSS) data. The NSS conducts every 5 years the Consumption expenditure survey (among other surveys) whence district specific per capita consumption expenditures are calculated. The last four NSS thick rounds <sup>36</sup> were undertaken in 1987-88 (43rd), 1993-94 (50th),

---

<sup>35</sup>The first three sources have been used by Iyer (2009)

<sup>36</sup>only the thick rounds can be used since the thin rounds have too few observations

1999-00 (55th), 2004-05 (61st). Given that district specific data is used, the 50th round is unusable since there are no district identifiers. For the purposes of this paper we use the other three rounds. The per capita consumption expenditure is calculated for all the districts using the Marginal per capita consumption (MPC hereafter), using the 30 day recall period since the 30 day recall period is the only one that is common across rounds. The MPC is used as a proxy for per capita income.

**Demography:** Apart from the Sample surveys by the NSS, India conducts a country wide census every 10 years. The data on demography and public goods are stored in the so called village directory data and is available for each census year since 1961. The 2001 census is used, whence we have data on total population, Schedule tribe and Caste population, population by religious groups. All the above information is available at the village level which has been used to construct the corresponding district level numbers.

**Geography:** We control for geographical terrain by the fraction of the districts uncultivated area that is barren rocky, sandy and steep sloping. We also control for remoteness by the log of the distance from the state capital. Such data is made available in the Wasteland atlas of India, Department of Land Resources (Ministry of Rural development) in collaboration with the National Remote Sensing agency, (Department of Space). These variables by definition do not vary over time. We also have data on forest cover in the different districts. The data comes from the various "State of forest cover" reports of the Forest Survey of India (FSI) and gives the percentage of forest cover in each district. The 2005 data is used.

**Land distribution:** The data on operational holdings of agricultural land comes from the Agricultural census of India which collects such data every 5 years. The data gives the number of operated landholdings in the different size classes (viz. Small, semi medium, medium and, large). Both the 1991 and the 2001 censuses are available. The gini coefficient for land inequality is calculated using this data. The computed gini coefficient varies from 0.12 to 0.78 and 0.14 to 0.79 in 1991 and 2001 respectively. The average inequality has gone up from 0.47 in 1991 to 0.5 in the year 2001 (data on Bihar not available in 2001). The land inequalities across the 2 years are highly correlated with Correlation coefficient of 0.97. We use the 1991 data since that way we have data prior to the high conflict years and also no major state is missing.

**Colonial Land Institutions:** This data comes from Banerjee and Iyer (2005). This data is available only for 233 districts, i.e. for districts which were directly under British control.

## B Figures

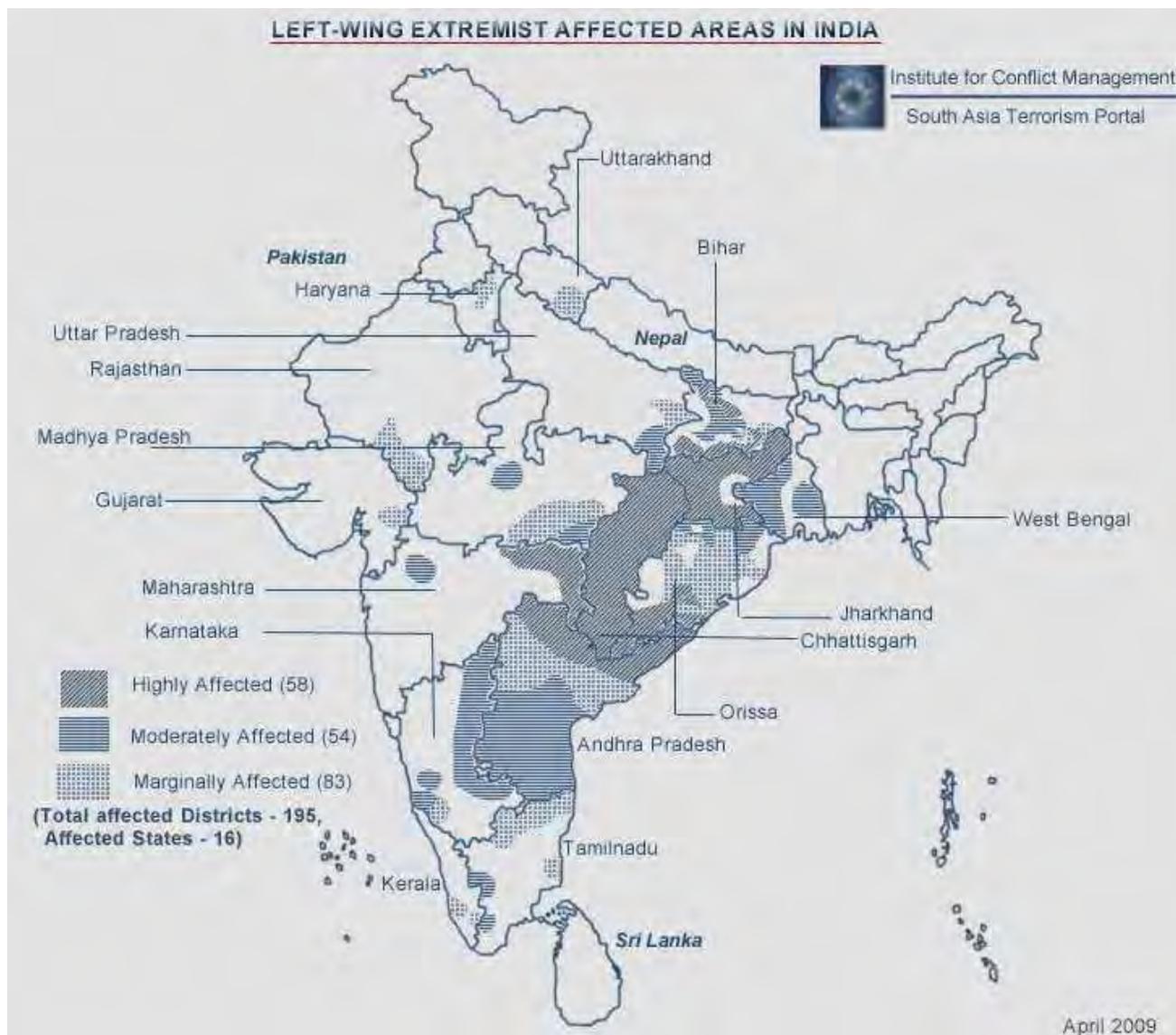


Figure B.1: Naxalite affected areas in India

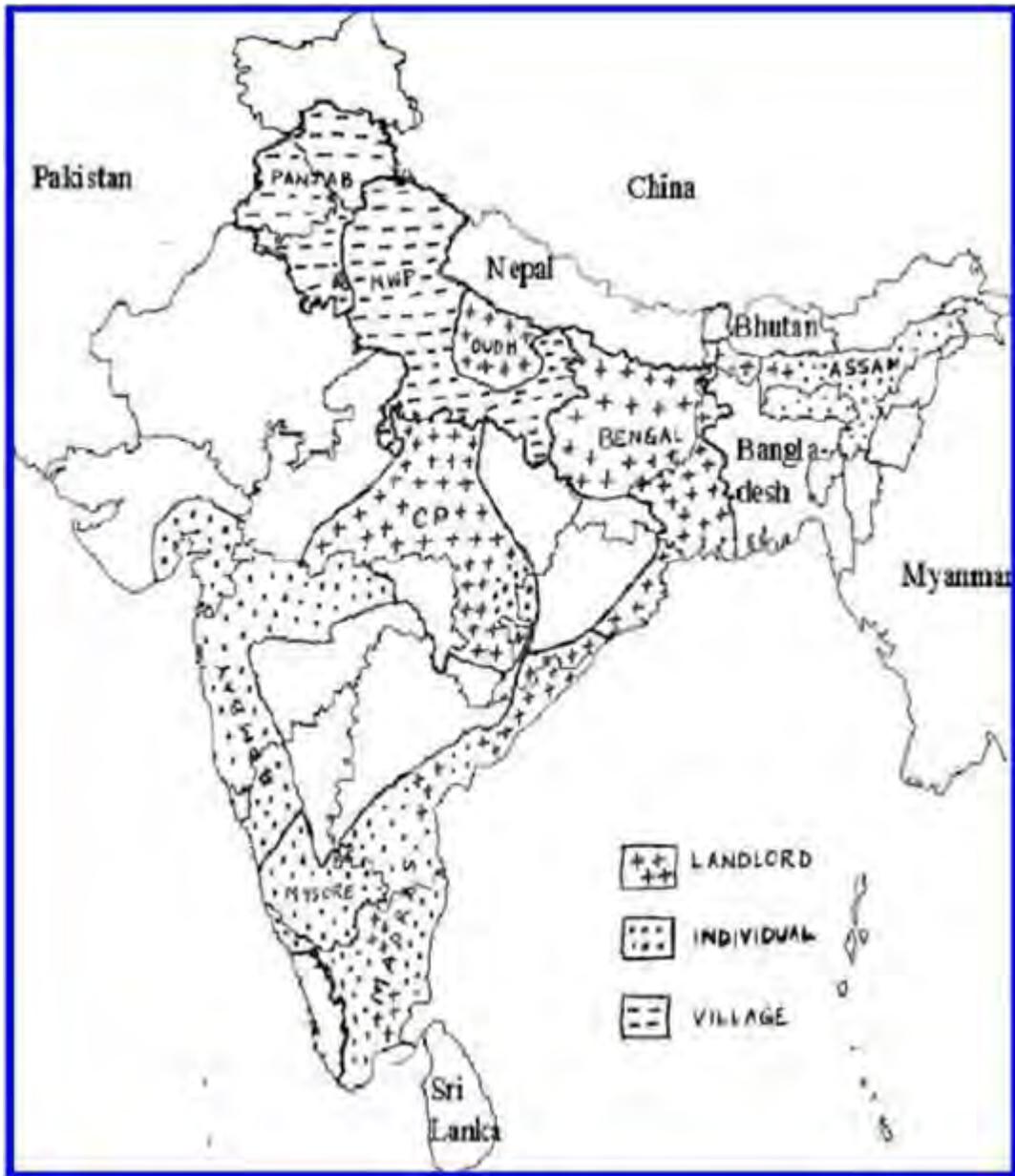


Figure B.2: Landlord & non Landlord districts in India (Source: Banerjee & Iyer, AER 2005)

## C Tables

Table C.1: Summary: All 3 rounds

Variable	Obs	Mean	Std. Dev.	Min	Max
maoist	1085	0.170507	0.376251	0	1
deadwounded	1085	6.322581	59.3128	0	1837
nincidents	1085	2.41106	15.79039	0	428
maoist_1	1085	0.087558	0.282781	0	1
gini	1046	0.266298	0.060809	0.103525	0.525915
mpc_87	1030	5.052897	0.244983	4.410083	5.777959
General_MPC_87	1030	5.132978	0.249794	4.498148	6.014436
SCST_MPC_87	1028	4.870675	0.250028	4.180412	5.941787
landineq91	1055	0.47359	0.178855	0.120449	0.788384
p_forest	1082	0.167094	0.18037	0.000699	0.832942
population density	1082	0.006383	0.022805	0.000692	0.416596
log_mpc	1046	5.849775	0.628196	4.410083	7.352093
log_gen_mpc	1043	6.002583	0.696857	4.498148	8.501844
log_scst_mpc	1045	5.674889	0.628981	4.180412	7.374076
log area	1082	8.721257	0.696394	6.475433	10.7288
proportion sandy	1040	0.006336	0.040049	0	0.688315
proportion barrenrocky	1040	0.00799	0.020191	0	0.265584
proportion steepaloping	1040	0.002556	0.010052	0	0.129534
log state capital distance	1031	5.485296	0.811263	0	6.899219
SCST_percent	1082	0.292082	0.15533	0.016113	0.946497
log_mpc_lag	686	5.608404	0.620172	4.410083	6.848523
log_gen_mpc_lag	685	5.727645	0.666179	4.498148	7.336274
log_scst_mpc_lag	685	5.440889	0.634257	4.180412	7.374076
p_nland	698	0.522766	0.429809	0	1

Table C.2: Correlation Matrix

	maoist	maoist_l	log mpc	proportion sandy	log state capital distance	proportion barren rocky	proportion steep sloping	proportion forest	SC percent	ST percent	population density
maoist	1										
maoist_l	0.5449	1									
log mpc	-0.2713	-0.0596	1								
prop sandy	0.1275	0.0804	0.1252	1							
log state capital distance	0.1327	0.1216	-0.0443	-0.1017	1						
proportion barren rocky	0.0462	0.0597	0.1443	-0.0774	0.0162	1					
proportion steep sloping	-0.1567	-0.0983	0.3246	-0.0626	-0.0396	0.1753	1				
proportion forest	0.114	0.0574	0.0694	-0.1653	0.2297	-0.0088	0.2107	1			
SC percent	-0.0293	-0.0003	0.1513	0.1606	-0.4011	-0.1188	0.0439	-0.2466	1		
ST percent	0.0197	-0.0132	-0.2652	-0.1673	0.1996	0.2404	0.0084	0.3179	-0.507	1	
population density	0.0463	0.0192	-0.0216	0.2251	-0.3502	-0.4447	-0.2079	-0.4327	0.1952	-0.5248	1
log area	0.2169	0.1613	-0.2474	-0.2124	0.2529	0.1663	-0.1171	0.1413	-0.1718	0.3598	-0.5955
gini	0.0218	0.1937	0.4101	-0.0043	0.119	-0.0208	0.0338	0.1363	-0.0309	0.0506	-0.0973
land inequality	0.1982	0.1453	0.0285	0.1137	-0.1999	-0.0849	0.0174	-0.0965	0.169	-0.437	0.661
MPC 87	-0.1784	-0.0812	0.6891	0.0768	-0.1229	0.1648	0.3058	0.0201	0.2413	-0.198	-0.0544
prop no landlord	-0.35	-0.2043	0.4941	-0.1214	0.1642	0.1929	0.3112	0.1289	-0.0381	0.0082	-0.3227
log general MPC	-0.2192	-0.0029	0.7489	0.0151	-0.0204	0.1275	0.1908	0.154	0.108	-0.0567	-0.0953
log SC MPC	-0.247	-0.0699	0.8392	0.1908	-0.0318	0.1126	0.3283	0.1087	0.0935	-0.0788	-0.0874
log ST MPC	-0.2613	-0.027	0.6032	0.0468	-0.1384	0.1341	0.2576	-0.0831	0.1407	-0.2957	0.1105
General MPC growth	-0.1843	-0.1657	-0.0077	-0.0175	0.0385	0.0135	-0.0405	0.1162	0.0213	-0.0096	-0.0512
SC MPC growth	-0.1742	-0.1985	-0.0731	0.0554	0.0077	0.0356	-0.0382	-0.1022	-0.0173	-0.0148	-0.0049
ST MPC growth	-0.1482	-0.0598	-0.0384	-0.0101	-0.1007	-0.0254	-0.0358	-0.1512	0.073	-0.1668	0.1571
MPC growth	-0.1861	-0.2043	-0.0069	0.05	0.0083	0.005	-0.0052	-0.0522	0.0139	-0.098	0.0441
	log area	gini	land inequality	MPC 87	prop non landlord	log general MPC	log SC MPC	log ST MPC	General MPC growth	SC MPC growth	ST MPC growth
log area	1										
gini	0.1235	1									
land inequality	-0.4394	-0.077	1								
MPC 87	-0.2575	0.101	-0.0619	1							
prop no landlord	0.1053	0.1455	-0.253	0.4601	1						
log general MPC	-0.1714	0.4456	-0.0005	0.5304	0.3819	1					
log SC MPC	-0.1468	0.3167	-0.0384	0.5509	0.3794	0.5422	1				
log ST MPC	-0.3048	0.1431	0.1097	0.4364	0.2944	0.4418	0.4632	1			
General MPC growth	-0.0727	-0.3791	0.0565	-0.0156	-0.0025	0.2722	-0.0521	-0.0138	1		
SC MPC growth	0.0245	-0.525	-0.0146	-0.0818	-0.0073	-0.2372	0.1186	-0.0525	0.5474	1	
ST MPC growth	-0.1332	-0.4111	0.1093	-0.0482	-0.0132	-0.1229	-0.0685	0.4045	0.4145	0.5549	1
Mpc growth	-0.0554	-0.5205	0.0658	-0.0829	-0.0028	-0.1254	0.0142	-0.0078	0.7043	0.8869	0.6367

Table C.3: Same sample as Table 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
consumption per capita	-1.282*** (0.434)	-1.371*** (0.500)	-1.143 (0.822)	-2.956** (1.234)	-2.144** (1.034)	-3.625 (2.812)
land inequality	2.743*** (0.631)	2.537*** (0.875)	4.552*** (0.863)	9.854*** (1.971)	9.773*** (2.086)	14.32*** (2.550)
proportion sandy	19.20*** (7.300)	25.04*** (6.797)	30.47** (13.77)	38.95 (70.41)	-1.637 (59.24)	43.93 (59.49)
log state capital distance	0.127 (0.197)	0.155 (0.213)	0.139 (0.210)	0.199 (0.483)	0.220 (0.352)	-0.0377 (0.338)
proportion barrenrocky	17.07*** (3.299)	24.39*** (4.083)	20.34*** (5.448)	46.46*** (16.97)	51.86*** (13.40)	37.59*** (9.129)
proportion steepslping	-70.86** (33.45)	-113.6*** (42.47)	-155.9*** (44.47)	-201.2*** (60.66)	-331.1*** (116.6)	-389.2*** (95.12)
proportion forest cover	1.211* (0.651)	1.817*** (0.686)	2.645*** (0.563)	5.579*** (2.041)	6.019*** (1.960)	7.412*** (1.737)
%Scheduled Castes	-0.862 (1.565)	-0.386 (1.704)	4.110*** (1.416)	-2.788 (4.282)	-3.855 (3.971)	11.19*** (3.395)
%Scheduled Tribes	4.071 (2.475)	5.503** (2.588)	1.813 (1.288)	1.400 (6.599)	3.974 (8.201)	0.939 (2.815)
%Scheduled Tribes square	-6.590 (4.192)	-9.555** (3.910)	-1.620 (2.259)	-5.864 (9.987)	-14.50 (13.57)	1.763 (3.586)
population density	0.536** (0.208)	0.789** (0.329)	0.401 (0.309)	1.327** (0.551)	0.836 (0.526)	0.291 (0.626)
log area	0.768*** (0.198)	0.870*** (0.289)	0.894*** (0.144)	3.006*** (0.544)	2.669*** (0.478)	2.698*** (0.338)
income inequality	-3.148* (1.780)	-5.625*** (1.848)	-0.423 (1.794)	-15.45** (6.189)	-18.13*** (6.455)	-2.350 (6.539)
initial consumption per capita		-0.116 (0.797)			-3.275*** (0.945)	
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	655	431	655	655	431	655
ll	-186.0	-138.6	-142.8	-651.3	-540.6	-592.8

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

In this table we have the same sample as in Table 2 but without the institutions variable. The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Table C.4: The role of income differences across ethnic groups: SC&amp;ST vs. General Castes

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
General Caste consumption pc	-0.492* (0.279)	-0.566** (0.229)	-0.722*** (0.181)	-2.169*** (0.575)	-1.418** (0.721)	-1.808*** (0.553)
SC & ST consumption pc	-1.037*** (0.389)	-1.039** (0.502)	-0.625 (0.547)	-2.222** (0.927)	-2.099* (1.165)	-2.449* (1.472)
land inequality	2.837*** (1.084)	2.902** (1.218)	4.486*** (1.004)	14.51*** (3.438)	12.97*** (3.905)	14.77*** (3.438)
proportion sandy	-1.091 (2.636)	-0.771 (2.417)	-0.0382 (1.379)	-46.54 (28.97)	-49.09** (24.44)	-15.83 (11.55)
log state capital distance	-0.0494 (0.124)	0.0640 (0.118)	-0.0234 (0.200)	-0.610*** (0.166)	-0.309* (0.184)	-0.346 (0.249)
proportion barrenrocky	10.27*** (2.192)	14.06*** (3.016)	11.88*** (4.389)	30.11** (14.30)	38.16** (15.75)	23.97*** (5.205)
proportion steepsloping	-65.18*** (19.23)	-90.58*** (35.12)	-125.6*** (34.34)	-238.9*** (40.98)	-206.4*** (52.41)	-324.7*** (75.77)
proportion forest cover	1.388* (0.718)	1.823** (0.783)	2.300*** (0.548)	2.452 (1.875)	2.511 (1.697)	5.491*** (1.526)
%Scheduled Castes & Tribes	0.982 (0.608)	1.704** (0.734)	1.976*** (0.330)	3.192* (1.849)	3.604** (1.614)	4.314*** (1.331)
population density	0.340 (0.222)	0.625** (0.299)	0.329 (0.286)	-0.472 (0.622)	-0.235 (0.544)	-0.140 (0.491)
log area	0.696*** (0.216)	0.784*** (0.304)	0.703*** (0.144)	2.357*** (0.387)	2.335*** (0.322)	2.504*** (0.291)
income inequality	-0.581 (1.937)	-4.248** (1.874)	1.396 (1.281)	-4.110 (3.588)	-10.37** (4.289)	-3.628 (3.982)
Initial consumption per capita	No	Yes	No	No	Yes	No
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	993	651	993	993	651	993
ll	-272.4	-180.9	-193.8	-1013.9	-825.3	-912.2

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

In this table we combine the incomes of the Scheduled Castes and tribes together and compare its impact on conflict vis a vis the income of general castes. The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Table C.5: The role of income differences across ethnic groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
General Castes Consumption pc	-0.474* (0.245)	-0.570** (0.253)	-0.702*** (0.211)	-0.234 (0.978)	-0.575 (0.973)	-1.021 (0.965)
Scheduled Castes Consumption pc	-0.463 (0.371)	-0.943** (0.417)	-0.370 (0.389)	-1.123 (0.908)	-1.605* (0.874)	-1.912** (0.945)
Scheduled Tribes Consumption pc	-0.747** (0.311)	-0.539** (0.267)	-0.664* (0.388)	-1.319* (0.739)	-0.867* (0.525)	-1.726* (1.038)
proportion sandy	16.08** (6.513)	25.89*** (4.803)	27.67** (11.69)	-40.88 (28.14)	-62.09*** (23.65)	2.766 (25.29)
log state capital distance	0.262 (0.233)	0.266 (0.259)	0.194 (0.259)	0.152 (0.570)	0.308 (0.408)	-0.0104 (0.465)
proportion barrenrocky	22.12*** (4.907)	26.58*** (5.385)	26.58*** (8.323)	57.78*** (18.87)	45.52*** (11.36)	43.53*** (12.02)
proportion steepslping	-110.4*** (33.05)	-81.10** (40.31)	-199.9*** (66.74)	-364.8*** (129.3)	-313.7*** (92.62)	-519.4*** (168.1)
proportion forest cover	1.657*** (0.559)	2.248*** (0.719)	2.808*** (0.757)	5.460*** (1.282)	4.763*** (1.303)	7.177*** (1.739)
%Scheduled Castes	-0.0800 (1.607)	0.531 (1.757)	4.503*** (1.599)	-2.296 (3.912)	-3.157 (3.723)	8.422*** (3.202)
%Scheduled Tribes	0.305 (1.999)	2.021 (2.027)	0.352 (1.304)	-5.308 (3.827)	-5.231 (4.501)	-2.864 (3.940)
%Scheduled Tribes square	-2.050 (3.880)	-5.541 (3.674)	0.640 (2.411)	0.270 (7.383)	-1.772 (9.236)	4.277 (5.165)
population density	0.282 (0.248)	0.616* (0.364)	0.260 (0.359)	0.0281 (0.720)	-0.0317 (0.769)	-0.392 (0.612)
log area	0.836*** (0.226)	1.060*** (0.363)	0.834*** (0.162)	3.482*** (0.535)	3.329*** (0.452)	2.772*** (0.434)
income inequality	-0.247 (1.626)	-2.970* (1.693)	1.169 (1.572)	-15.78*** (5.502)	-18.36*** (5.540)	-5.041 (6.186)
land inequality	3.062*** (0.459)	2.605*** (0.742)	4.118*** (0.901)	12.64*** (2.648)	10.91*** (2.363)	14.10*** (2.595)
proportion Non landlord	-0.973*** (0.254)	-1.041*** (0.263)	-1.216** (0.533)	-2.377*** (0.644)	-1.982*** (0.503)	-1.023 (0.828)
initial consumption per capita		0.298 (0.601)			-0.741 (0.614)	
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	507	319	507	507	319	507
ll	-139.9	-101.8	-113.7	-535.3	-440.7	-498.7

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

This table is same as Table 4 in the paper but also controls for historical institutions. The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Table C.6: The Effect of Growth on Conflict

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability	Probability	Probability	Intensity	Intensity	Intensity
Consumption growth	-0.00951 (0.0178)	-0.0102 (0.0163)	-0.00426 (0.0192)	0.0213 (0.0416)	0.00485 (0.0446)	-0.0304 (0.0601)
proportion sandy	21.75*** (6.644)	19.16*** (6.258)	49.08** (24.75)	-6.925 (45.93)	-21.67 (32.90)	5.918 (26.62)
log state capital distance	0.211 (0.213)	0.247 (0.216)	0.243 (0.249)	-0.204 (0.356)	-0.0299 (0.335)	0.258 (0.552)
proportion barrenrocky	22.84*** (5.202)	25.71*** (4.944)	26.42*** (7.329)	41.98** (18.18)	50.00*** (16.85)	36.16*** (12.15)
proportion steepsloping	-47.14* (28.57)	-100.8*** (36.85)	-105.4** (50.28)	-137.2*** (53.07)	-291.2** (120.0)	-203.7*** (62.36)
proportion forest cover	2.138*** (0.720)	2.158*** (0.673)	3.158*** (0.979)	3.942* (2.211)	5.177** (2.169)	7.002*** (1.942)
%Scheduled Castes	-0.614 (1.319)	-0.609 (1.453)	4.886*** (1.837)	-6.801* (3.815)	-7.115 (4.564)	6.799** (3.162)
%Scheduled Tribes	4.870** (2.386)	4.791* (2.721)	4.472** (2.019)	-0.980 (6.152)	-0.794 (6.837)	0.641 (5.397)
%Scheduled Tribes square	-8.931** (3.735)	-8.838** (3.945)	-6.013** (2.740)	-6.719 (11.11)	-9.050 (12.37)	-0.467 (6.949)
population density	0.846*** (0.297)	0.806** (0.334)	0.734* (0.444)	-0.367 (1.021)	-0.181 (0.776)	1.175 (0.978)
log area	1.059*** (0.269)	1.059*** (0.322)	1.217*** (0.213)	3.041*** (0.579)	3.107*** (0.483)	3.312*** (0.558)
income inequality	-7.272*** (1.467)	-7.164*** (1.574)	-4.630*** (1.503)	-26.04*** (5.552)	-23.95*** (6.529)	-12.25** (5.949)
land inequality	2.013** (0.833)	2.216** (0.893)	3.041** (1.285)	9.880*** (3.227)	10.24*** (2.894)	6.677*** (1.629)
proportion Non landlord	-1.162*** (0.290)	-1.114*** (0.244)	-1.552*** (0.427)	-2.702*** (0.572)	-2.298*** (0.543)	-2.191*** (0.650)
initial consumption per capita		-0.151 (0.731)			-2.409*** (0.616)	
State dummies	No	No	Yes	No	No	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	422	420	422	422	420	422
ll	-135.4	-132.9	-105.6	-544.4	-535.4	-500.7

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

This table is same as Table 5 in the paper but also controls for historical institutions. The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.

Table C.7: Change in causes of conflict over time by subgroups

	Round 1 Probability	Round 2 Probability	Round 3 Probability	Round 1 Intensity	Round 2 Intensity	Round 3 Intensity
General Caste consumption pc	0.863 (1.418)	-1.620*** (0.534)	-0.464 (0.501)	-0.246 (4.885)	-2.258 (1.486)	-0.607 (0.858)
Scheduled Caste consumption pc	0.322 (0.870)	-1.302* (0.685)	-0.459 (0.882)	3.526 (4.910)	-0.370 (0.887)	-1.219 (0.996)
Scheduled Tribe consumption pc	-0.730 (0.534)	-0.487 (0.706)	-1.028 (1.216)	-2.752* (1.539)	-1.881 (1.671)	-1.694 (1.567)
land inequality	5.158*** (1.296)	8.944*** (1.701)	5.965*** (1.419)	21.03*** (3.429)	26.51*** (6.472)	15.70*** (3.588)
%Scheduled Tribes	1.691 (3.532)	-3.502 (2.235)	4.303 (4.225)	3.225 (10.65)	1.630 (5.616)	6.614 (4.066)
%Scheduled Tribes square	-3.811 (6.525)	6.782*** (2.241)	-3.609 (4.572)	-7.120 (13.58)	1.148 (5.354)	-11.35* (5.930)
%Scheduled Castes	-2.530 (2.727)	-0.871 (1.377)	5.477*** (1.856)	-6.067 (8.469)	-8.409 (6.048)	-5.112 (4.129)
proportion sandy	-21.10 (25.46)	-3.417 (6.103)	-1.458 (5.286)	-130.7** (64.37)	-327.9*** (81.03)	-57.31*** (16.78)
proportion barrenrocky	1.241 (4.836)	20.94** (8.550)	16.39** (8.266)	52.56 (44.10)	22.67 (25.40)	36.79* (20.62)
proportion steepsloning	-119.2*** (42.79)	-334.7** (132.8)	-64.11* (34.31)	-709.7*** (240.5)	-407.0** (194.0)	-280.5*** (82.66)
log state capital distance	-0.291 (0.180)	0.139 (0.339)	0.199 (0.256)	-0.934*** (0.297)	-0.911*** (0.282)	-0.108 (0.282)
proportion forest cover	-0.174 (1.346)	2.228** (1.053)	1.979 (1.873)	2.856 (2.781)	-3.910** (1.606)	1.147 (1.937)
log area	1.043*** (0.375)	0.899** (0.444)	1.376*** (0.493)	6.405*** (1.855)	3.707*** (0.826)	4.014*** (0.451)
population density	-0.518 (0.569)	-0.358 (0.529)	1.037** (0.466)	0.859 (1.596)	-2.691 (1.688)	0.389 (0.508)
income inequality	4.769 (3.376)	-1.449 (4.559)	-3.852 (4.060)	-0.849 (13.98)	-8.148 (11.81)	-14.10*** (5.003)
Conflict_1	Yes	Yes	Yes	Yes	Yes	Yes
Observations	289	136	145	289	152	145
ll	-58.81	-39.90	-33.24	-144.8	-215.9	-329.0

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients and not marginal effects. The data is from a database built by the author combining myriad databases. All regressions at the district level with robust standard errors, clustered at the state level (in parentheses). Columns 1,2 and 3 explain the probability of conflict (the presence of conflict in the district) using Probit regressions. Columns 4, 5 and 6 explain the intensity of conflict (the no. of dead & wounded in the district) using Negative Binomial regressions. There are 3 time periods used corresponding to 3 NSS rounds, 1987-88 (43rd), 1999-00 (55th), 2004-05 (61st). The conflict data is for the years 1979-2009, which are clubbed to corresponding NSS rounds.