Civil conflict with rising wages and increasing state capacity: Theory and application to the Maoist insurgency in India

Anand Shrivastava*

3rd November 2016

Abstract

This paper presents a model demonstrating the mechanisms through which rising wages and increasing state capacity could lead to civil conflict over a natural resource. Conflict is modelled as a game between the state, the rebels and the local population. Increasing wages lead to lower risk aversion, which makes the local population more likely to support conflict whose outcome is uncertain. The relationship between conflict and state capacity is non-monotonic because while the rebel's willingness to fight decreases with increasing state capacity, the state's willingness to fight increases with it. The implication regarding wage is tested on district level data on conflict and agricultural wage from six Indian states that are affected by the Maoist insurgency. The implementation of the National Rural Employment Guarantee Act is used as an instrument for wage after removing district specific heterogeneity. The results support the model's implication and are robust to different specifications.

JEL: D74, O13

Keywords: Civil conflict, natural resources, India, NREGA, Maoist insurgency

*Azim Premji University, Bangalore, India. (e-mail: anand.shrivastava@apu.edu.in). I am grateful to Toke Aidt, Jean-Marie Baland, Sandeep Baliga, Tim Besley, Steven Durlauf, Julien Gagnon, Alex Harris, Sanjeev Goyal, Sriya Iyer, Renate Hartwig, Hamish Low, Gerard Padro i Miquel, seminar participants at University of Cambridge, IEA World Congress 2014, Barcelona Summer Forum, CEPR/PODER summer school at University of Cape Town, Doctoral Workshop at Université de Namur and Université catholique de Louvain, for helpful comments and suggestions. I would also like to thank Joseph Gomes and Kristian Hoelscher for being generous in sharing data. I am grateful to the Gates Cambridge Foundation for financial support. This paper is funded under the grant "Policy Design and Evaluation Research in Developing Countries" Initial Training Network (PODER), which is funded under the Marie Curie Actions of the EU's Seventh Framework Programme (Contract Number: 608109).

1 Introduction

'What causes civil conflict?' is a widely studied question in political economy. One of the most important causes described in the literature is low wages. Early economic models of conflict (Hirshleifer, 1988; Grossman, 1991; Skaperdas, 1992) have agents optimizing between production and forceful appropriation. In these models, production is the opportunity cost of fighting and hence the probability of conflict decreases with increasing returns from production, i.e. wages. Besley and Persson (2011) present a model of political violence, which includes civil war and repression, as a two period game between an incumbent and an opposition. Here too they find that higher wages would lead to lower likelihood of conflict as higher wages increase the opportunity cost of building an army. Most empirical studies on civil conflict look at the impact of income, instead of wages, and find that lower incomes are correlated with higher probability of conflict (Collier and Hoeffler, 2004; Fearon and Laitin, 2003; Miguel et al., 2004; Hegre and Sambanis, 2006). Another important cause of civil conflict described in the literature is low state capacity (Fearon and Laitin, 2003; Buhaug, 2006). A state with fewer financial resources, often measured either as GDP or as tax revenue, will have lower repressive capacity, hence the cost of rebelling will be low leading to higher insurgency and conflict. Bates (2008) provides empirical evidence on how weaker governments in Africa might have led to increased conflict.

Much of the empirical literature, however, consists of cross-country analyses that use variables derived from national income measures as the explanatory variable. Among the few exceptions is the paper by Berman et al. (2011) which looks at the relationship between unemployment and conflict in Afghanistan, Iraq and the Philippines and rejects the predictions of the opportunity cost explanation as it does not find a positive correlation between unemployment and conflict. The case of India, where conflict with Maoist rebels has been spreading and intensifying over the last decade while the economy of the country has been doing well, also seems to contradict the logic provided by prevailing models.

The Maoist conflict in India has its roots in left-wing extremist movements started in the 1960s but has increased in intensity and frequency of incidents in the last decade, with 3000 conflict related deaths recorded between 2008 and 2011. Figure 1 shows two scatter plots of conflict versus wages using district level panel data from six Indian states. The vertical axis has residuals after removing state, time and state-time interacted fixed effects from the number of incidents of conflict. The first plot shows the full sample, while the second one is after removing outliers. In both cases there is a significant positive correlation, instead of the expected negative correlation. The positive correlation becomes even more important as the effect of conflict on wages, which confounds most empirical work in this area, is likely to be negative. Figure 2 shows trends in state capacity, measured as

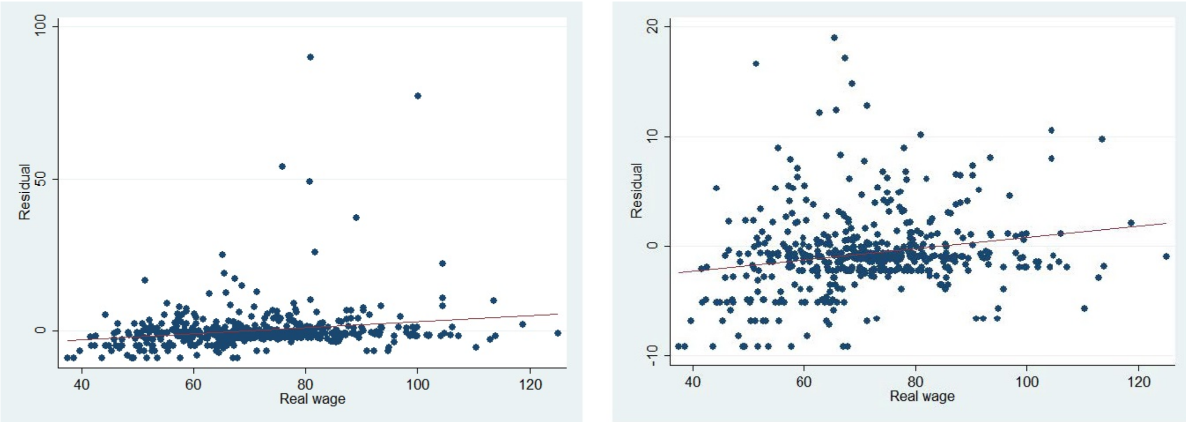

Figure 1: Scatter plot of incidents of conflict versus real wage at district level
Source: Gomes (2015), South Asian Terrorism Portal and Agricultural Wages of India, Ministry of Agriculture.

the total revenue expenditure in six affected states, and conflict measured as number of incidences of Maoist violence in these states. The state capacity has increased by almost 50% in this duration, but this does not seem to have reduced conflict as predicted by the literature.

This paper presents a model demonstrating the mechanisms through which rising wages and increasing state capacity could increase the likelihood of civil conflict. The model is a strategic game between the state, the rebels and the local population, with the conflict occurring over control of a natural resource. The local population derives its livelihood from this resource, which the state wants to take control of. The rebels oppose the state for ideological reasons, but need the support of the local population to fight. Conflict happens when the state deploys forces to capture the resource, and the local population supports the rebels in fighting to win it back.

Comparative statics on the model show that at low wage levels conflict becomes more likely with increasing wages. The intuition behind this can be explained as follows. Conflict is risky as its outcome is uncertain. The decision for the local population to support the rebels and enter into a conflict is a trade-off between the cost of conflict, in terms of destruction of property and chances of arrest or death, and the uncertain benefit of winning control of the resource with some probability. It can be thought of as a lottery where they can win the prize of the natural resource with some probability. Whether buying the lottery is optimal or not depends, among other things, on the risk aversion in the local population. At low wages people will be more risk averse and are hence less likely to prefer buying the lottery of conflict. As wages increase, people become less risk averse and conflict becomes more likely. This does not exclude the opportunity cost effect explained earlier which makes conflict less attractive as wages increase. But at low wage

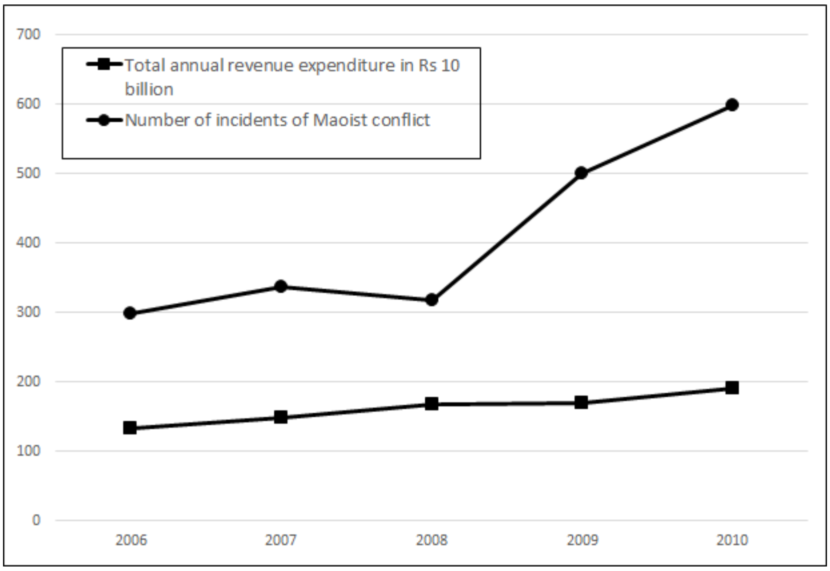

Figure 2: Trends in state capacity and conflict
Source:Gomes (2015), South Asian Terrorism Portal and Reserve Bank of India.

levels, the risk aversion effect dominates the opportunity cost effect, leading to a positive relationship between wages and conflict.

The other main result of the model is that the effect of state capacity on conflict is non-monotonic with the two being positively correlated under certain circumstances. As state capacity increases, it does make conflict more costly for the rebels, but a more powerful state is also more likely to try and capture natural resources thus opening up new fronts of conflict. Hence, the state's willingness to fight increases with state capacity while the rebels' willingness to fight decreases with state capacity and this generates the non-monotonicity between state capacity and conflict.

In the second part of the paper, the prediction of the model regarding the effect of wages on conflict is tested on data from the Maoist conflict in central India. District level data on incidents of Maoist conflict from six affected states for the period 2006-10 is used as the dependent variable, which is regressed on real wage for agricultural labour as the explanatory variable, along with a number of control variables including forest cover, tribal population, urbanisation and presence of mining industries. It is found that the coefficient on real wage remains positive for a number of different specifications and is statistically significant for most of them. Sources of potential bias are addressed using internal instruments from the panel data, as well as an external instrument, the implementation of the National Rural Employment Guarantee Act. We find that one standard deviation change in the real wage, which is Rs 14.5, can cause two additional incidents of violence in the average district. The positive correlation between wages and conflict is robust to different estimation methods and sub-samples and thus supports the model's prediction.

This paper contributes to the literature in three ways. First, it uses a game-theoretic model to explain how increasing wages can lead to civil conflict by lowering risk aversion. Most economic models of conflict (Hirshleifer, 1988; Grossman, 1991; Skaperdas, 1992; Besley and Persson, 2011) do not take into account the role of risk aversion and predict that increasing wages would lower the likelihood of conflict because of increased opportunity cost of fighting. The only other model that has a similar prediction to the one presented here is by Dal Bó and Dal Bó (2011) who use a general equilibrium model of the economy to show that positive income shocks may increase the risk of conflict if they happen in the capital intensive sector while they have the opposite effect in the labour intensive sector. They treat conflict more as crime rather than political conflict and model it accordingly as the third sector of the economy, after the capital and labour intensive ones. In contrast, the model presented here is that of a political conflict between the state and rebels and the effect of changes in the wages of the local population is analysed.

Second, the model shows that increasing state capacity will not always reduce the likelihood of conflict as the relationship between state capacity and conflict is non-monotonic. The discussions about the effect of state capacity on conflict in the literature have been varied and nuanced but these have mostly been about what state capacity actually means (Fjelde and De Soysa, 2009; Hendrix, 2010). This paper used a simple definition of state capacity as the total financial resources available to the state that it can potentially deploy as repressive forces. This is close to the original definition of state capacity as “the power of the state to raise revenue” (Besley and Persson, 2010). Using this definition, state capacity is incorporated into an economic model of conflict and a novel prediction about its effect on conflict is made.

Third, the empirical analysis on the Maoist conflict in central India shows that increasing wages have led to an increase in conflict. This is in contrast to several cross-country studies that have found that higher incomes lead to lower conflict (Collier and Hoeffler, 2004; Fearon and Laitin, 2003; Miguel et al., 2004; Hegre and Sambanis, 2006). Even most within-country studies on economic causes of Maoist insurgencies, Do and Iyer (2010) on Nepal, and Gomes (2015) and Hoelscher et al. (2012) on India, find that poorer regions are more likely to have conflict. All these studies use some measure of per capita income rather than wages as their explanatory variable. By using agricultural wage, which is likely to be the wage of the people involved in the conflict, this paper provides a cleaner test of the predictions regarding the effect of wage on conflict. Two papers find results similar to the one presented here. The first is Khanna and Zimmermann (2014) who also look at the Maoist insurgency in India and find that the implementation of a government employment guarantee program, the National Rural Employment Guarantee

Act in India, may have increased conflict. The second is Berman et al. (2011) who find a negative correlation between unemployment and conflict in Afghanistan, Iraq and the Philippines.

This paper has important policy implications. The Indian government's response to the increase in Maoist violence has been, in accordance with the prevailing logic, along two lines - development and repression. The government has increased development activities in conflict hit areas while also increasing the repressive capacity of the states that are affected by insurgency (Planning Commission of India, 2008; Ramesh, 2012). This has had limited success in reducing the insurgency. The theoretical and empirical evidence presented in this paper indicates that when the reason for conflict is the absence of rights for people who are directly dependent on natural resources for their livelihood, then neither development nor repression may solve the problem, at least in the short run. The solution may lie in solving the underlying issue of rights. Legislation to this affect has already been passed in the form of the Forest Rights Act (Ministry of Law and Justice, Government of India, 2006) and enforcing this legislation may be a more effective way of containing the insurgency.

The following section describes a game-theoretic model of civil conflict and presents comparative statics based on the model. Section II presents some background and related literature on the Maoist conflict in India. Section III describes the data used. Section IV presents the results of the empirical analysis. Section V concludes.

2 Model

Most economic models of conflict analyse conflict at the level of the nation state, where a rebel group fights an incumbent group to take control of the state's resources in the form of tax revenues or natural resource rents. This model is of conflict at the sub-national level. We shall refer to the area under consideration as a locality, which could be a village or a group of villages that utilize and value a common natural resource. The existence of a rebel group outside the locality is assumed. Most developing countries have such armed rebel groups whose existence may be more because of ideological, religious or ethnic opposition to the government rather than any economic reasons. The formation of these rebel groups is not modeled here and it is assumed that they had been formed at some point in history and are ready to enter the locality. The model presented here looks at the occurrence of armed conflict between the rebels and the state in the locality under consideration. Another way to look at it is as a model of the spread of conflict to a hitherto peaceful locality. The reward for winning the conflict is control of this particular locality and hence of the natural resource, not of the entire state.

The local population is the third party in the conflict between the state and the rebels. Support from sections of the local population who are disaffected with the state is crucial to rebellion, but has generally been neglected in conflict models. Here it is assumed that these people, who are poor and disadvantaged, care not so much about the ideological motives of the rebels, but about economic issues of livelihood and sustenance. They are modeled as deriving their livelihoods from a natural resource, for example a forest, over which they have no property rights, and from wage labour. The state can decide to capture the resource for industrial use, for example mining, thus making this section of the population worse off and making it more likely that they will support the rebels. If the rebels win then these people get back access to the natural resource, i.e. the rebels do not have the capacity or the will to exploit the resource. The potential commitment problem between the rebels and the people is resolved by assuming that the rebels would want to honour their commitment as a way to get further support from people in other localities in a repeated setting. The industrial use of the resource requires significant investment that can happen only under the state and the rebels are not in a position to undertake the industrial exploitation of the resource, thus reducing their incentive to break their commitment to the locals.

It is in this setup that the effect of changes in wages and in state capacity on conflict is analysed. The conflict is modeled as a game between three players - the state, the rebels and the locals. The resource is valued by two sections of the society - the locals, let us call them section T as they would use the resource through traditional means, and section M, who would use the resource through modern industrial means. The state is controlled by M and its objective function is expected utility of M.[1] In the first stage of the game, the state decides to divide its budget between expenditure on security forces that can give it control of the natural resource and on public goods that benefit both T and M. The total budget available to the state is considered a measure of state capacity. In the second stage, T decides whether to support the rebels or not, by comparing its expected utility in conflict with that when there is no conflict . In the third stage, the rebels decide on the amount of forces they are going to deploy in the locality. They make this decision based on an exogenously given value that they place on taking control of the locality. If T doesn't support the rebels, then the chances of the rebels winning are zero.[2] This represents the scenario when the rebels use guerrilla tactics and rely on the locals for shelter, provisions or at the very least, not letting the state know of their locations. If T does support the

[1]This assumption can be relaxed with the state's objective function being a weighted average of the expected utilities of T and M. This will not change the results obtained from the model.

[2]This assumption too can be relaxed by assuming that if T does not support the rebels then their chances of winning would be greatly reduced but not necessarily become zero. Again, the main results of the model will not change but the algebra will become much more complicated.

rebels, then given the forces deployed by the state and the rebels, the probability of one side winning the conflict is given by a contest function. If the state wins the conflict, then M gets control of the natural resource. If the rebels win the conflict then T gets control of the resource.

2.1 Setup

The population in a locality consists of two sections. Section T consists of households depending on wage labour and traditional means of livelihood. Section M consists of those who earn their livelihoods through modern and industrial means.

A representative household in T derives utility from private consumption c and from a public good G provided by the state.

$$u_T(c,G) = g_T(c) + f(G)$$

We make the following assumption about the utility function

A1. $g'_T(.), f'(.) > 0; g''_T(.), f''(.) < 0$ and $g'_T(c) \to \infty$ as $c \to 0$.

This says that the utility is increasing in consumption and in public goods but at a decreasing rate and that as consumption reaches zero, the marginal utility of consumption increases without bound. We are also assuming additive separability between utility from consumption and that from public goods.

Let consumption be given by

$$c = c(w, \beta, S)$$

$w \geq 0$ is wage

$\beta \in \{0,1\}$ denotes whether T has access to the resource ($\beta = 1$) or not ($\beta = 0$)

$S \in \{0,1\}$ denote whether T supports ($S = 1$) the rebels or not ($S = 0$)

We make some assumptions abut the functional form of $c(.)$

A2. c and $\dfrac{\partial c}{\partial w}$ are finite for all finite values of w, β and S.

A3. $\dfrac{\partial c(w,0,1)}{\partial w} > 0$

A4. $c(w,\beta,0) > c(w,\beta,1)$

A5. There exists $w_c > 0$ such that $c(w_c,0,1) = 0$

A2 and A3 are self-explanatory. A4 says that there is always a cost of supporting the rebels that leads to a reduction in consumption. A5 implies that there exists a positive wage level, where the cost of supporting the rebels would reduce consumption to zero.

We make one more assumption, which, for a given functional form, puts restrictions on the range of w for which the analysis is valid.

A6 $c(w,1,1) > c(w,0,0)$

This implies that the benefit of getting access to the resource is larger than the cost of supporting the rebels. If this is not valid then we can trivially see that there will be no conflict.

Let us take an example. Let $g_T(c) = c^{1/\gamma}$. Let a be the value of the resource, h_1 be the fraction of wage lost as opportunity cost of supporting the rebels and h_2 be the fixed cost of supporting the rebels. Consumption is given by:

$$c(w,\beta,S) = \beta a + w - S(h_1 w + h_2)$$

We can check that all the assumptions will be satisfied for $\gamma > 1,\ 0 \leq h_1 < 1,\ h_2 > 0$ and $0 < w < (a - h_2)/h_1$. We will return to this example later in the analysis.

For notational simplicity let us denote $u_T(c(w,\beta,S),G)$ as $u_T(w,\beta,S,G)$.

Since access to the resource (β) depends on the result of the conflict, T's expected utility is given by

$$U_T = P\, u_T(w,1,S,G) + (1-P)u_T(w,0,S,G)$$

Where P is the probability of the rebels winning the conflict and is given by

$$P = P(F,SR)$$

Where F is the spending by the state on security forces and R is the resources committed by the rebels for fighting. We make three sets of assumptions about the conflict technology.

A7. $P(F,0) = 0$, for $F > 0$, and $P(0,0) = 1$

A8. $\dfrac{\partial P}{\partial F} < 0,\ \dfrac{\partial^2 P}{\partial F^2} > 0,\ \dfrac{\partial P}{\partial SR} > 0,\ \dfrac{\partial^2 P}{\partial SR^2} < 0$

A9. $\dfrac{\dfrac{\partial^2 P}{\partial F^2}\dfrac{\partial P}{\partial SR}}{\dfrac{\partial P}{\partial F}} < \dfrac{\partial^2 P}{\partial F \partial SR} < \dfrac{\dfrac{\partial P}{\partial F}\dfrac{\partial^2 P}{\partial SR^2}}{\dfrac{\partial P}{\partial SR}}$

A7 says that if the state deploys a positive amount of forces and the rebels don't deploy any forces or they don't get support from T, then the state is sure to win. The second part of A7 defines the default such that when there is no conflict, T controls the resource. A8 says that there are positive but decreasing returns to fighting for both sides. A9 puts limits on the strategic complementarity in the conflict technology. These assumptions are analogous to the assumptions made in Besley and Persson (2011).

The state's objective function is section M's utility, which is given by

$$u_M(1-\beta,G) = g_M(1-\beta) + f(G)$$

Hence, the state's expected utility is

$$U_M = P(F,SR)g_M(0) + (1-P(F,SR))g_M(1) + f(G)$$

Without loss of generality, we can set $g_M(0) = 0$ and let $g_M(1) = k$.

$$U_M = (1 - P(F, SR))\, k + f(G)$$

The state can spend its revenues on public goods G, or on security forces F. e is the amount available to the state for expenditure,which can be seen as a measure of state capacity. Since $f'(G) > 0$, the budget constraint will bind. Hence, we can replace G, with $e - F$. The state's problem is:

$$\max_{F \in [0,e]} U_M = (1 - P(F, SR))\, k + f(e - F)$$

The objective function of the rebels is

$$U_R = P(F, SR)v - R$$

Where v is the value of winning the conflict in this locality and will depend on the strategic importance of the locality for the rebels and the amount of extortion revenue that the locality can generate.

We do not consider the effect of wages on the recruitment of soldiers or insurgents by the states or the rebels, as modeled in Grossman (1991). One reason for this is that as we are looking at conflict at a local level, the insurgents and soldiers involved may not have been recruited locally. Even if their is some recruitment at the local level, it will just be another form of labour that offers the prevailing wage w, and will not change the utility function of a household in T. The conflict function P itself is modeled in terms of the resources spent, assuming that any changes in wage will affect recruitment, if any, on both sides in the same manner, leaving the probabilities unchanged.

2.2 Strategic interaction

The strategic interaction is modeled as a three stage game. The timing is

1. The state chooses security forces $F \in [0, e]$.
2. T chooses $S \in \{1, 0\}$ to decide whether or not to support the rebels
3. The rebels allocate $R \geq 0$ for fighting.
4. The payoffs are realised.

The state moves slowly, so it is plausible that it will move first and anticipate the reactions of the other players. The sequence of T and the rebel's moves is not backed by any strong reasoning. If the sequence is reversed we get the same kinds of equilibria with slightly different equilibrium conditions, but the main conclusions remain unchanged. If they move simultaneously, what we have essentially is a co-ordination game with a peaceful and a conflict equilibrium. In this formulation, the players have no way of choosing the equilibrium that is better for them. In practice, there are a number of ways to solve the coordination problem, but our intention is not to model those. Hence, we choose sequential moves to allow the 'better' equilibrium to be chosen and choose the sequence which allows easy interpretation.

We characterise the equilibrium for given values of the exogenous variables *w, k, v* and *e.* Our equilibrium concept is subgame perfect Nash equilibrium and we solve backwards.

In stage 3, the rebel's problem is

$$\underset{R\geq 0}{max} U_R = P(F,SR)v - R$$

If the rebels are supported by T then $S=1$ and the interior solution $\hat{R}(F)$ is given by

$$\frac{\partial P(F,\hat{R})}{\partial R} = \frac{1}{v}$$

If they are not supported by T then $S=0$. For $S=0$ and $F>0$, $U_R = -R$ and for $S=0$ and $F=0$, $U_R = v-R$. In either case, the optimal response is $R=0$.

Therefore, the rebel's equilibrium strategy is

$$R^*(F,S) = S\hat{R}(F)$$

This means that the rebels will allocate forces only if supported by *T*. The optimal amount of forces if supported, $\hat{R}$, will depend on the value v that the rebels put on controlling the locality and on the security forces F deployed by the state in stage 1.

In stage 2, T's problem is

$$\underset{S\in\{0,1\}}{max} U_T = P(F,SR^*(F,S))u(w,1,S,G) + (1-P(F,SR^*(F,S))u(w,0,S,G)$$

Substituting $R^*(F,S) = S\hat{R}(F)$ and using $S^2 = S$, we get

$$\underset{S\in\{0,1\}}{max} U_T = P(F,S\hat{R}(F))u(w,1,S,G) + (1-P(F,S\hat{R}(F)))u(w,0,S,G)$$

It is evident that if $F=0$, then $S=0$ is the best response, as it would mean $U_T = u(w,1,0,G)$, which is the highest possible value (using A4 and A6). This means that if the state does not deploy any forces to capture the resource then T does not have any reason to support the rebels. If $F>0$, then T will compare two possible scenarios.

$$U_T\mid_{S=0} = u(w,0,0,G)$$

$$U_T\mid_{S=1} = P(F,\hat{R}(F))u(w,1,1,G) + (1-P(F,\hat{R}(F)))u(w,0,1,G)$$

We now define critical probability P_c such that $P(F,\hat{R}(F)) > P_c \Rightarrow U_T\mid_{S=1} > U_T\mid_{S=0}$, i.e. if the probability of the rebels winning is greater than the critical probability P_c, then it is optimal for T to support the rebels. P_c is given by

$$u_T(w,0,0,G) = P_c.u_T(w,1,1,G) + (1-P_c).u_T(w,0,1,G)$$

$$Or, \qquad P_c = \frac{u_T(w,0,0,G) - u_T(w,0,1,G)}{u_T(w,1,1,G) - u_T(w,0,1,G)} = \frac{g_T(c(w,0,0)) - g_T(c(w,0,1))}{g_T(c(w,1,1)) - g_T(c(w,0,1))}$$

P_c only depends on w and using A5 and A6, we can see that $P_c \in [0,1]$. A lower value of P_c means that it is more likely that the probability of the rebels winning, $P(F,\hat{R}(F))$, will be greater than P_c, which means that T is more likely to support the rebels. Hence, a lower value of P_c implies a higher likelihood of conflict.

Lemma 1 : *There exists a positive threshold level wage $\tilde{w}$ such that for wages less than $\tilde{w}$, P_c decreases with an increase in wage, i.e.* $\dfrac{\partial P_c}{\partial w} < 0$.

This is shown mathematically in Appendix I, but the intuition is as follows. Because of the concavity of the utility function, at low wage levels the risk aversion, as measured by, say, the Arrow-Pratt coefficient of absolute risk aversion, is very high. Hence, for any given probability of winning the conflict, it is more likely at low wages that T will find conflict too risky and will choose to not support the rebels. As wages increase from this low level, the risk aversion decreases and for the same probability of winning, T may now find it optimal to support the rebels and have an conflict because of lower risk aversion. So, as wages increase T requires a lower probability of winning to support the rebels and start a conflict and thus we get the inverse relation between P_c and w.

There are two other ways in which P_c is affected by wages. First, the incremental utility of getting access to the resource decreases as wages increase. Second, if there is an opportunity cost of conflict, this increases with wages. Both these relationships would make conflict less likely with increase in wages and hence lead to a positive correlation between P_c and w. But at low wages the first relationship mentioned above is much stronger than the other relationship, since the marginal utility is higher at lower levels of consumption, thus giving us the required result.

Let us return to the example mentioned earlier.

$$g(c(w,\beta,S)) = [\beta a + w - S(h_1 w + h_2)]^{1/\gamma}$$

Figure 3 shows how P_c varies with w for this functional form for the values $\gamma = 2$, $a = 100$, $h_1 = 0.1$, $h_2 = 10$.

The graph shows a U-shaped relationship between P_c and w. In Lemma 1 we have proved that the downward sloping part of the curve will always exist, thus ensuring that as we increase wages from very low levels, there will always be a shift from no conflict to conflict.

Proposition 1: *Given values of v and P_c, there exists a value of security forces F_c such that it is optimal for T to choose $S = 0$ for all $F \geq F_c$; and for $w < \tilde{w}$ this F_c increases with increasing wage i.e.* $\dfrac{\partial F_c}{\partial w} > 0$.

This means that if the state manages to deploy security forces to a level F_c, then the probability P of the rebels winning, even when they deploy the optimal level of forces given F_c, will be so low that T will prefer not to support the rebels. As wage increases and T becomes less risk averse, it may find that at this value of P conflict is now optimal.

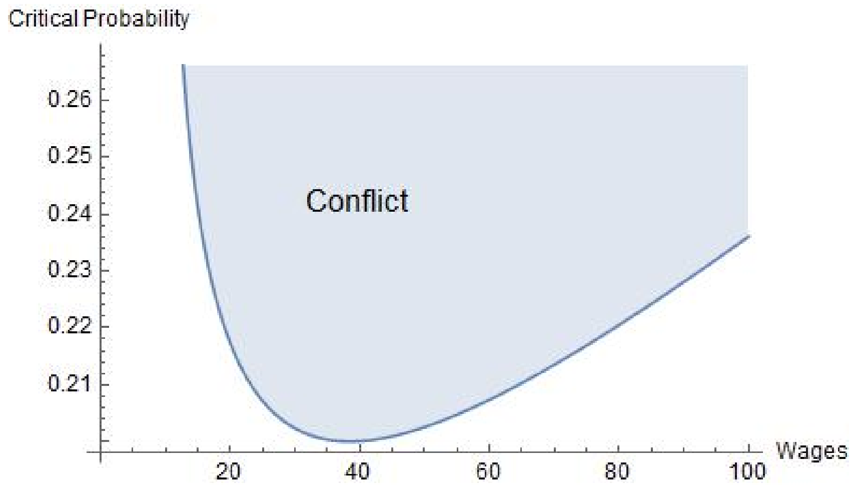

Figure 3: Graph of critical probability of winning (P_c) versus wages (w)

Hence, the state will need to lower P even more by deploying even more forces to make conflict not optimal for T again. Hence, we get a positive relationship between F_c and w.

The proof is given in Appendix II and it is shown that F_c is given by

$$P(F_c, \hat{R}(F_c)) = P_c$$

If the state increases F beyond this level, then the probability of the rebels winning will go below P_c, thus making it optimal for T to not support the rebels. As P_c decreases, the state will require more force to bring the probability of the rebels winning below P_c. Hence, F_c and P_c will have an inverse relationship. This taken along with Lemma 1 gives us the second part of the proposition. Note that as P_c depends only on w, F_c will depend only on w, the wage, and v, the value the rebels put on controlling the locality.

Thus, T's equilibrium strategy is

$$S^*(F) = \begin{cases} 1 & if\ 0 < F < F_c \\ 0 & otherwise \end{cases}$$

In stage 1, the state's problem is

$$\max_{0 \le F \le e} U_M = (1 - P(F, S^*(F)R^*(F, S^*(F)))) \, k + f(e - F)$$

We show in Appendix III that the equilibrium strategy for the state is

$$F^* = \begin{cases} 0 & if\ k \le min\left\{ f(e) - f(e - F_c), \dfrac{f(e) - f(e - \hat{F})}{1 - P(\hat{F}, \hat{R}(\hat{F}))} \right\} \\ \hat{F} & if\ \dfrac{f(e) - f(e - \hat{F})}{1 - P(\hat{F}, \hat{R}(\hat{F}))} \le k \le \dfrac{f(e - \hat{F}) - f(e - F_c)}{P(\hat{F}, \hat{R}(\hat{F}))} \\ F_c & if\ k \ge max\left\{ f(e) - f(e - F_c), \dfrac{f(e - \hat{F}) - f(e - F_c)}{P(\hat{F}, \hat{R}(\hat{F}))} \right\} \end{cases}$$

The equilibrium strategy of the state shows a trade-off between the value k of the resource and the utility of public goods. If the resource is much less valuable than the public goods, then the state will not want to sacrifice any public goods to acquire it and will not deploy any forces. If the resource is much more valuable than the public goods, then the state will deploy as much force as is required to get the resource for sure i.e. $F = F_c$. In the intermediate case, the state will sacrifice some public goods to get the resource with the probability $(1-P)$.

The subgame perfect Nash equilibrium is given by the equilibrium strategies of the three players: $F^*, S^*(F)$ and $R^*(F,S)$. The equilibrium will be one of the following types depending on the values of w, e, k and v.

1. Peace: $F = 0; S = 0; R = 0$

$$\text{if} \quad k \leq min\left\{ f(e) - f(e-F_c), \frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \right\}$$

2. Conflict: $F = \hat{F}; S = 1; R = \hat{R}(\hat{F})$

$$\text{if} \quad \frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \leq k \leq \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))}$$

3. Repression: $F = F_c; S = 0; R = 0$

$$\text{if} \quad k \geq max\left\{ f(e) - f(e-F_c), \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))} \right\}$$

Note that the values of F_c, $\hat{F}$ and $\hat{R}(\hat{F})$ can be determined given the values of w, e, k and v.

2.3 Comparative statics

We want to conduct comparative statics on two margins. One is between the equilibria of Peace and Conflict and the other between Repression and Conflict. The margin between Peace and Repression is not considered here as it will not be observable in conflict data. We look at the effect of state capacity and wages on these two margins.

2.3.1 Repression-Conflict margin

We can use the equilibrium conditions to get the following expression for the Repression-Conflict margin:

$$RC = P(\hat{F},\hat{R}(\hat{F}))k - [f(e-\hat{F}) - f(e-F_c)]$$

Repression will result if $RC \geq 0$ and Conflict if $RC \leq 0$. This is a trade-off for the state between the natural resource and public goods as moving from Conflict to Repression increases the probability of getting k from $(1-P)$ to 1, at the cost of reducing public goods provision from $(e-\hat{F})$ to $(e-F_c)$.

Recall that as wage increases from a low level, T's risk aversion decreases making it more likely to support the rebels. Hence the state needs to deploy more forces to sufficiently reduce the probability of the rebels winning to make it optimal for T to not support the rebels. Hence at low wage levels, $\frac{\partial F_c}{\partial w} > 0$. In the expression above, as wage increases F_c will increase making Repression more costly, thus reducing RC and leading to Conflict. This gives us our main theoretical result that as wage increases from a low level, conflict becomes more likely.

The effect of state capacity is analysed in detail in Appendix IV, but intuitively we can see that increased state capacity makes it easier for the state to increase security forces to F_c hence making Repression more likely than Conflict.

2.3.2 Peace-Conflict margin

The Peace-Conflict margin is given by the following expression:

$$PC = [f(e) - f(e - \hat{F})] - (1 - P(\hat{F}, \hat{R}(\hat{F})))k$$

Peace will result if $PC \geq 0$ and Conflict if $PC \leq 0$. Conflict increases the probability of getting k from 0 to $(1 - P)$, at the cost of reducing public goods provision from e to $(e - \hat{F})$ compared to Peace.

Since there is no F_c in this expression, wage does not have any effect on this margin. The effect of state capacity is analysed in Appendix IV, but intuitively it can be explained as follows. The state may choose $F = 0$, if the probability of winning the conflict is low and is not worth the cost in terms of the loss of public goods. Now if more revenue becomes available to the state, the probability of winning the conflict equilibrium may be high enough for the state to choose conflict. Hence, increased state capacity may lead to Conflict. A further increase in state capacity may lead to the state having enough resources to allocate F_c and hence lead to the Repression as described earlier. This second relationship is the one that has been highlighted in literature when describing a negative correlation between state capacity and conflict.

From the analysis above we obtain two potentially testable implications

1. At low wage levels, conflict is more likely with an increase in wages w.

2. The effect of state capacity, e, on conflict is non-monotonic. At the Peace-Conflict margin, the likelihood of conflict increases with state capacity, but at the Repression-Conflict margin, it decreases with an increase in state capacity.

The first implication about the relationship between wages and conflict is a novel contribution to the literature. As discussed earlier, most of the economic literature on civil conflict predicts an inverse relation between wages and conflict (Grossman, 1991; Besley and

Persson, 2011). In the implication about state capacity, only the effect at the Repression-Conflict margin is taken into account in most studies (Fearon and Laitin, 2003). The opposite effect at the Peace-Conflict margin has not been modeled in the literature. It is possible to potentially test for both these implications, but getting reliable measures of state-capacity at the local-government level is difficult. Here we are only going to test the first implication using data from the Maoist insurgency in India.

3 The Maoist insurgency in India

3.1 Background

During the 1960s, a number of armed left-extremist movements started in parts of India, primarily in the states of West Bengal, Bihar and Andhra Pradesh. These groups were influenced by the Maoist philosophy of armed rebellion and were part of fundamentally communist movements of peasants rebelling against landlords and fighting for land redistribution. Over time the movements split into several groups and did not pose a substantial threat to the state although sporadic acts of violence continued till the late 1990s. From the year 2000, groups started becoming more active in the tribal districts of India mostly in the states of Jharkhand, Chattisgarh and Odisha. Many of the different insurgent groups combined to form the Communist Party of India (Maoist) in 2004 and the frequency of incidents has grown steadily since then. In the four years from 2006-2009, forty new districts were affected by Maoist violence.[3] The Maoist propaganda has also changed from class warfare to 'jal, jungle, jameen' -water, forest and land (Guha, 2007). As of 2011, around 183 districts[4] had experienced some form of violence and 83 districts were designated as seriously affected by Maoist violence (Government of India, 2011). Over 10,000 people have been killed in this conflict (Hoelscher et al., 2012) with over 3000 deaths during 2008-11 (Government of India, 2011).The threat of Maoist violence has increased to such a level that former Prime Minister Manmohan Singh had called it 'the single biggest internal security threat facing the country'.

This spread of violence has happened during a period of increasing economic prosperity in India. Wages as well as state revenue have increased throughout the country. Increased wages should increase the opportunity cost of conflict for the rebels, and increased state revenue should increase the cost of rebellion as the state has the increased repressive capacity, both resulting in lower conflict. Hence, the increase of violence goes against the intuition provided by standard conflict models.

[3]Calculated based on the data obtained from Gomes (2015).

[4]Districts are administrative units in India comprising of towns and villages. India has 640 districts with an average population of 1.8 million people per district.

A number of theories have been promulgated to explain this surge in violence and the spread of the Maoists into the tribal areas. Many view the Maoists as terrorists and the insurgency as a problem of terrorist violence (Simeon, 2010). There are also those who view this as a result of the tribals being left out of the growth experienced by the rest of the economy (Guha, 2007). Many scholars have also noted the fact that the areas populated by tribals overlap with the mineral rich areas of the country, which are also covered in dense forests. These tribals are from tribes like Ho, Santhal, Munda, Muria and Kond who have traditionally depended upon natural resources for their livelihood through activities like collection of honey, bamboo, leaves of tendu and sal etc. from the forest, collection of firewood from forest, collection of drinking water from water bodies and shifting cultivation in forest land. These natural resources also have alternative uses for mining, timber, hydro power projects or even for wildlife conservation. In general, the tribals don't have any formal rights over these natural resources. The state can and does use them for these alternative purposes either directly or by allotting them to the private sector. This has a very serious impact on the livelihoods of tribals and they are often driven to destitution. This might have provided a sympathetic audience and fertile recruitment ground for the Maoists, thus explaining the surge in violence in tribal districts. A number of instances of state 'expropriation' of such resources have been reported.[5] These facts allow us to apply our model to this situation and compare its predictions with the data.

The Indian Government's action against the Maoists has been mainly through Operation Greenhunt launched in 2009 where central paramilitary forces along with the state police are used to tackle the Maoists with a 'clear, hold, build' strategy which also ostensibly incorporates development initiatives. Quite separate from this the government has been pursuing a number of welfare measures including the National Rural Employment Guarantee Act (NREGA) to improve conditions in all poor areas of the country (Planning Commission of India, 2008) but also with a view to reducing 'left-wing extremism'. The model described here provides a framework in which we can think of the potential impact of these two approaches - using security forces and providing economic benefits - on the spread of the Maoist insurgency.

3.2 Economic literature on the Maoist insurgency in India

There have been a few recent economic studies of the Indian Maoist conflict. Hoelscher et al. (2012) use district GDP per capita as the income variable and establish a negative relationship with conflict using cross-sectional data. Gomes (2015) uses a panel with three time periods spread over 15 years. He finds a negative correlation between conflict

[5]For an overview of such instances see pp 221-227 of Shrivastava and Kothari (2012) and Padel and Das (2006).

and income implying that over a long period of time, poorer areas are more likely to have conflict compared to more well-off areas. Kapur et al. (2012) look for causal effects of shocks to natural resources in the form of vegetation on the intensity of conflict, using rainfall as an instrument. They find that twice lagged vegetation is negatively correlated with conflict in states with a significant population of tribals. Eynde (2011) analyses a particular aspect of the insurgency - the targets of Maoist violence. The paper investigates the factors that effect the decision of the Maoists to target security forces and informants respectively. Fetzer (2014) finds that the social insurance offered by the National Rural Employment Guarantee Act (NREGA) reduces the income elasticity of conflict.

The papers mentioned above either look at a different aspect of the conflict than the effect of wages (Kapur et al., 2012; Eynde, 2011; Fetzer, 2014), or look at the effect of income on conflict but either at a much longer time scale (Gomes, 2015) or using only cross-sectional data (Hoelscher et al., 2012).

The paper empirically closest to the results presented here is Khanna and Zimmermann (2014). They look at the effect of the National Rural Employment Guarantee Act on conflict using a regression discontinuity approach and find that it resulted in increased conflict. However, their explanation for their result is along the lines of Berman et al. (2011) implying that NREGA reduced the cost of information for the state and hence it was easier for the state to get information about the movements of the Maoist, thus leading to reduced conflict. The information channel is likely to result in the opposite effect as well, more information with the security forces could lead to more attacks on Maoist positions leading to more violence. We present some indicative evidence in the Results section to show that the effect of NREGA on conflict is likely to have been through wages.

4 Data

A panel for five years starting from 2006 covering all districts (total 148) of the six states of Andhra Pradesh, Bihar, Chattisgarh, Jharkhand, Odisha and West Bengal was constructed.[6] Over 95% of all Maoist related incidents from 2006-09 took place in these six states. The selection of these states is because the model assumes the existence of the rebel groups around the locality. This condition is satisfied in these states where Maoist activity was conspicuously present before the sample period begins.

The two important variables for testing implication 1 are conflict and wages.

Conflict is measured as the number of incidents of Maoist violence reported in a particular district. The model presented earlier analyses the likelihood of conflict in a 'loc-

[6]Districts that existed in 2005 were considered. Data from new districts formed later were merged with their parent districts

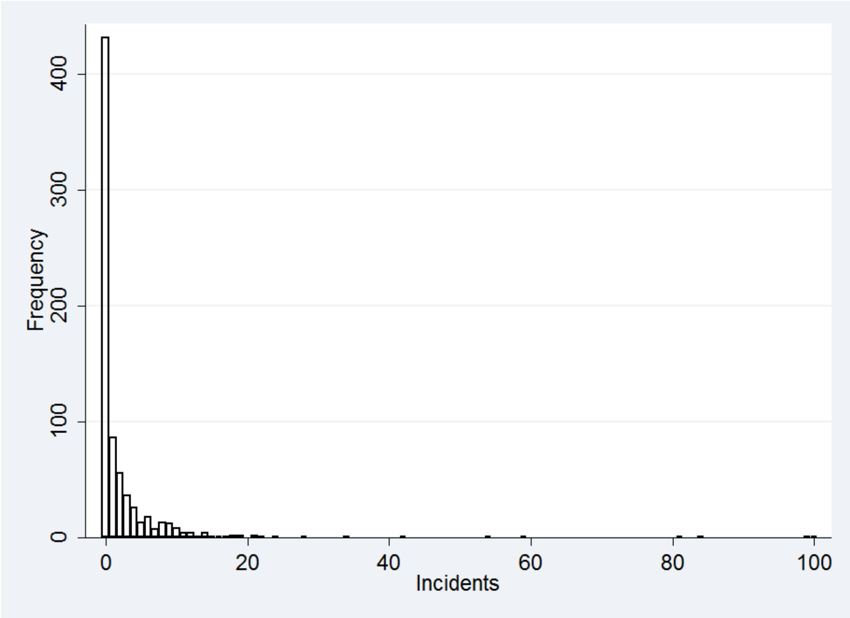

Figure 4: Histogram of annual number of incidents in a districtl

ality'. A district can be thought of as consisting of many such localities and an increased probability of conflict would mean that we would observe more incidents in that district because of conflict starting in more localities in the same district. In this respect, although we are looking at number of incidents in a given district, we can still interpret this as the extensive margin of conflict as depicted in the model, rather than the intensive margin. However, there is going to be some conflation between repeated incidents in the same locality and incidents in different localities in the same district that we are unable to differentiate here.

This data till 2009 is from Gomes (2015) and has been obtained mainly from the South Asian Terrorism Portal (SATP). The website collects data on the incidents from reports by news agencies and is hence able to provide district level data on number of incidents and the intensity of violence in terms of casualties. The data was extended till 2010 using SATP as wage data is available till that year. A histogram of the number of incidents is presented in Figure 4.

This is count data with a large number of zeroes. Hence, Poisson regression was considered a good starting point for the analysis. A binary control variable for the existence of Maoist activity before the analysis period is constructed using data on incidents in the period 2001-05. This variable is 1 for a district if any incident of Maoist related violence is reported from 2001-05 and 0 otherwise.

We use agricultural wages as a close proxy for the actual wages faced by those affected by state acquisition of natural resource. This is because agricultural is the primary occupation in rural India. The data on wages was collected from the Agricultural Wages of India (Ministry of Agriculture, Government of India, 6 09) reports. There are a number of other possible sources for data on rural wages. Himanshu (2005) presents a survey of all these sources and concludes that AWI is the best source. The closest contender is the National Sample Survey, but its surveys are not annual and the sample sizes vary substantially across rounds, hence leaving AWI as the most appropriate source. Other studies like Berg et al. (2013) and the World Bank study conducted by Özler et al. (1996) also use AWI data.

AWI reports provide daily wage rates for different agricultural trades (eg., ploughing, herding etc.) for each month in a year for each district for male, female and child workers. Sometimes data from multiple centres in a single district is reported. An annual average of the wages for male workers for each trade for each centre was obtained. Then a simple average of these values over all trades was taken for each centre. Ideally a weighted average of these should have been taken but data on number of people employed under each trade is not available. It is also probable that the same labourer might perform different tasks on different days. Berg et al. (2013) find high correlation between the wages for these tasks. In any case, rather than make any arbitrary judgement on including or excluding particular tasks or assigning weights, a simple average was considered a good approximation for the analysis. If there were multiple centres in a district then the average over these centres was taken. These values were corrected for inflation using the consumer price index for agricultural labourers for each state (Labour Bureau, Government of India, 2013). Hence we obtain a panel of real wages for agricultural labour.

The panel is unbalanced and has 524 observations instead of the total 740. Some states like West Bengal and Jharkhand do not have reports for certain years, while some districts just have missing data in particular years. This problem is tackled in greater detail in the sub-section on robustness checks. Another point to note is that the annual wage data is collected from July of one year to June of the following year. For the purpose of analysis, the annual average was considered to be for the second year, i.e. the wage for 2005-06 was considered to be the wage for 2006. A histogram of the real wage (2006 prices) is given in Figure 5. The average legal minimum wage at 2006 prices in these states over 2006-10 was around Rs 70/day.[7] A little less than half of all observations are below this level. Hence, we are dealing with relatively low wage levels here. Figure 6 shows the state-wise trend of average real wages for agricultural labour.

We use some control variables to account for other district level factors that can effect conflict. The main one is percentage forest cover as it can effect the conflict in multiple ways. A panel dataset for the forest cover was constructed using the biennial forest sur-

[7]Computed by the author based on data from Indiastat (www.indiastat.com) .

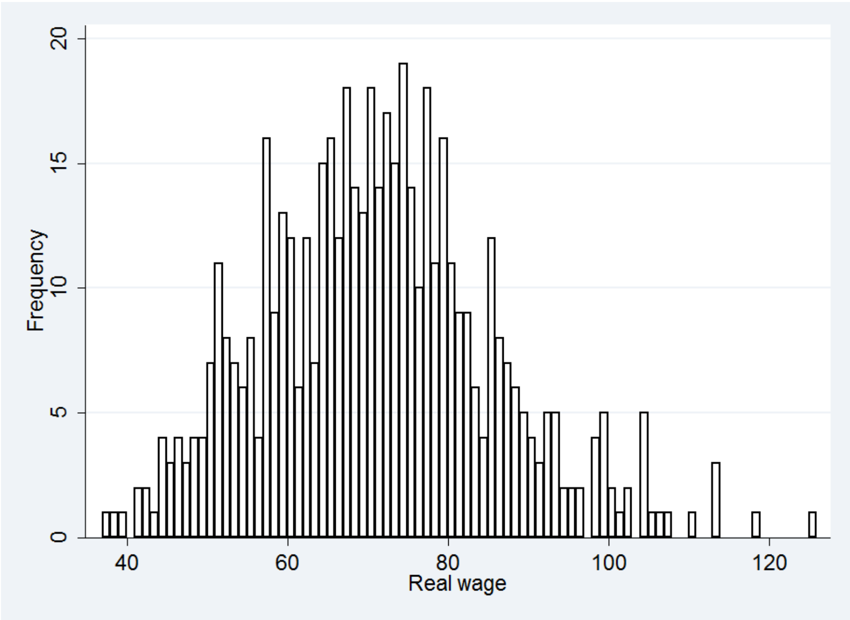

Figure 5: Histogram of district real wage at 2006 prices

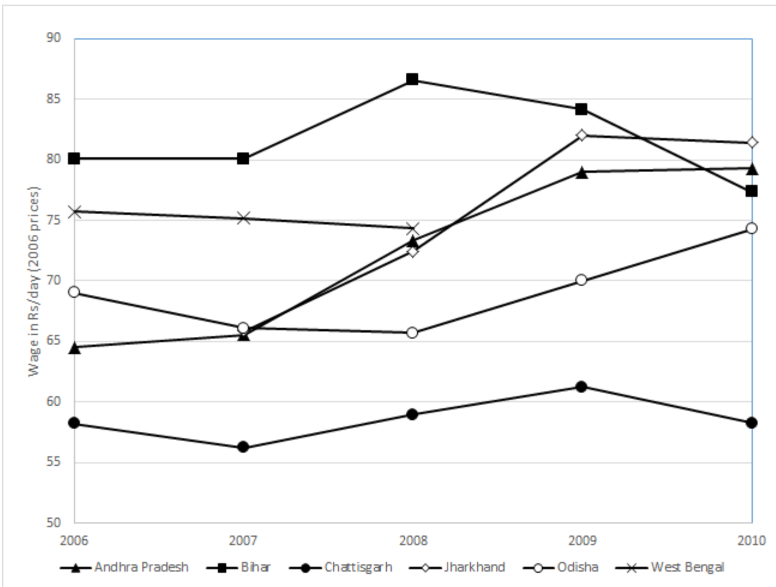

Figure 6: Trends in real wage for agricultural labour in the six states over 2006-10

veys conducted by the Ministry of Environment and Forests, Government of India (5 11). The other control variables used are - i) Share of scheduled tribe (ST) population in a district, ii) Whether the district has an active coal or iron mining operation, iii) Urbanisation. The data for these control variables has been taken from Hoelscher et al. (2012) and are cross-sectional data. The Scheduled tribe and urbanisation data are from the 2001 census (the data on 6 newly formed districts is missing) and the mining data is as of 2007. The list of all the variables used and their definitions is given in Table 1.

5 Results

Table 2 presents the results of regressions of *Number of incidents* on *Real wage*, with different specifications and estimation techniques. The standard errors are clustered at the district level as we are dealing with panel data and the error terms for the same cross-sectional unit may be correlated over different time periods. The first column of the top panel shows the results of an Ordinary Least Squares (OLS) regression where a number of control variables and the lagged dependent variable are included as regressors. We also include state-time fixed effects. This includes state fixed effects, time fixed effects and their interaction. This is important as the anti-insurgency operation is largely decided by the state governments, hence the difference in the government response to the Maoists will be reflected in time varying state fixed effect. This will also control for the time-varying propensity of state governments to acquire natural resources in tribal areas. The coefficient estimate for *Real wage* is positive and significant at 5%. The coefficient value of 0.025 indicates that an increase in the real wage by Rs. 40 can cause an additional incident of conflict

By using OLS, we are imposing a linear specification which may not be appropriate as *Number of incidents* is a count variable. Hence, in columns 2 and 3 of panel A, we use a Poisson specification to account for this. Column 2 uses Quasi Maximum Likelihood Estimation (QMLE) (Wooldridge, 2010). We also perform a one-step Generalised Method of Moments (GMM) estimation where the moment condition is only in terms of the conditional mean with a multiplicative error term (Windmeijer, 2006).[8] The results of the GMM estimation are in column 3. The coefficient estimate for *Real wage* is positive and significant at 5% for both the QMLE and the GMM regressions. The coefficient estimate here is a semi-elasticity. A coefficient of 0.014 means that an increase in real wage by 1 Rupee will increase the number of incidents by 1.4%. The average marginal effect is 0.038, this means that an increase of around Rs 26 in real wage is required to cause an

[8]The moment condition is $E\left(\frac{y_{it} - exp(x_{it}\beta)}{exp(x_{it}\beta)} \mid x_{it}\right) = 0$, where y_{it} is the number of incidents in district i in year t and x_{it} is the vector of explanatory variables.

Table 1: List of variables used

Name	**Type**	**Definition**	**Mean (S.D.)**	**Observations**
Number of incidents$_{it}$	Count	Number of incidents of Maoist violence or fighting between the Maoists and the security forces reported in district *i* in year *t*	2.54 (8.21)	740
Real wage$_{it}$	Continuous	The wage for agricultural labour in district *i* in a given agricultural year (July$_{t-1}$ - June$_t$) at 2006 prices	71.41 (14.49)	524
NREGA$_{it}$	Binary	Equals 1 in district i for all years $t \geq t^*$, where t^* is the year of implementation of the scheme in district *i*	0.51 (0.50)	740
Forest cover$_{it}$	Percentage	Percentage of the land area covered in forests	19.20 (17.22)	740
Maoist presence in 2001-05$_i$	Binary	Equals one if there was any Maoist incidents reported in district *i* during 2001-05, equals zero otherwise	0.53 (0.50)	740
Mining$_i$	Binary	Equals one if there was an iron ore or coal mine in district *i* in 2007	0.22 (0.42)	740
Percent scheduled tribes$_i$	Percentage	Percentage of the population belonging to one of the Scheduled Tribes according to the 2001 Census	15.87 (19.47)	740
Urbanisation$_i$	Percentage	Percentage of the total population that lived in urban areas according to the 2001 Census	16.61 (15.42)	740

additional incident.

A potential source of bias here is unobserved district specific heterogeneity. We can address this using fixed effects (FE) estimation at the district level. The results of a linear FE regression are presented in column 1 of panel B. The coefficient estimate for *Real wage* is positive but not significant, and it is slightly larger than earlier estimates but of the same order of magnitude. The cost of doing fixed effects estimation is that we lose the variation 'between' districts and are only relying on the variation 'within' districts. The 'within' standard deviation in *Real wage* is almost half of the 'between' standard deviation. This may be one of the reasons for the loss of significance. The linear 'between' regression, or regression using district means, is presented in column 4 of panel A and the coefficient of *Real wage* is an order of magnitude higher and significant at 1%.

We have also done regressions (not shown) with year fixed effects along with district fixed effects. It does not change any of the results but using year fixed effects leads to the instrument that we use becoming weak. Hence, we present results without year fixed effects for consistency.

By using fixed effect estimation we have taken care of the district specific unobserved heterogeneity, but we have introduced another bias resulting from the presence of the lagged dependent variable as a regressor, also known as Nickell bias. We can address this problem by using Arellano-Bond GMM estimation, where the regression equation is first differenced to remove the fixed effect, and then the twice lagged dependent variable is used as an instrument for the differenced lagged dependent variable (Arellano and Bond, 1991; Roodman, 2006). The results of the Arellano-Bond regression are shown in column 2 in the bottom panel. The coefficient estimate is similar to the one for linear FE and is significant at 5%. We manage to address the Nickell bias but we lose a lot of observations in the process because of the requirement of twice lagged dependent variable.

We do find that the actual coefficient estimate of the lagged number of incidents is statistically insignificant in the Arellano-Bond regressions. It may be possible that the lagged number of incidents was only representing the district level fixed effects in the pooled regressions, and on removing the fixed effects the effect of the lagged variable itself is not significant. Hence, we drop the lagged number of incidents for the remaining regressions. We conduct a linear FE regression again and find that the coefficient estimate of *Real wage* does not change much in magnitude or significance. This is shown in column 3. The FE and Arellano-Bond GMM that we have done so far have all assumed linearity. To address this we now conduct a fixed effects Poisson regression using conditional MLE (Hausman et al., 1984; Wooldridge, 1999). We find that the coefficient estimate for *Real wage* is positive but smaller than the pooled Poisson regressions conducted earlier, and is significant at 5%.

Table 2: Regressions of number of incidents on real wage and other controls

Dependent variable is *Number of incidents* $_{it}$				
Panel A				
	1	2	3	4
	OLS	Poisson QMLE	Poisson GMM	'Between' linear
Coefficients	ME	SE	SE	ME
Real wage$_{it}$ (Rs/day)	0.025 **	0.014 **	0.015 **	0.184 ***
	(0.011)	(0.007)	(0.008)	(0.051)
Number of incidents$_{it-1}$	0.752 ***	0.030 ***	0.086	
	(0.027)	(0.008)	(0.083)	
Forest cover$_{it}$	0.010	0.030 ***	0.049 ***	0.051
	(0.010)	(0.009)	(0.018)	(0.041)
Maoist presence in 2001-05$_{i}$	1.168 ***	1.538 ***	1.453 ***	4.185 ***
	(0.336)	(0.319)	(0.411)	(1.192)
Mining$_{i}$	- 0.563 *	- 0.770***	- 0.616	- 0.008
	(0.338)	(0.217)	(0.380)	(1.308)
Percent scheduled tribes$_{i}$	0.033 **	0.007	0.015 *	0.111 ***
	(0.013)	(0.007)	(0.009)	(0.040)
Urbanisation$_{i}$	0.005	0.008	0.015	- 0.022
	(0.012)	(0.007)	(0.011)	(0.044)
State-time fixed effects	Yes	Yes	Yes	State only
Number of observations	429	429	429	524
Number of districts	128	128	128	130
Panel B				
	1	2	3	4
	FE linear	Arellano-Bond	FE linear	FE Poisson
Coefficients	ME	ME	ME	SE
Real wage$_{it}$ (Rs/day)	0.030	0.038 **	0.030	0.010 **
	(0.020)	(0.017)	(0.019)	(0.005)
Number of incidents$_{it-1}$	0.078 *	0.097		
	(0.046)	(0.096)		
Forest cover$_{it}$	- 0.0002	0.002	- 0.053	- 0.016 ***
	(0.100)	(0.084)	(0.148)	(0.040)
Number of observations	429	300	524	377
Number of districts	128	120	130	90

ME denotes marginal effect and SE is semi-elasticity. In columns 1-3 of panel A and column 1 of panel B, we lost observations because of the lagged dependent variable. Column 2 in the panel B loses more observations as twice lagged dependent variable is used as instrument. Column 4 of panel B also loses observations because the fixed effects Poisson estimator omits districts where the dependent variable is 0 in all time periods. Column 4 of panel A and column 3 of panel B have all observations because of omitting the lagged dependent variable. Standard errors are clustered at the district level, except for the column 4 in Panel A, and are given in parentheses. Significance levels: *** - 1%. ** - 5%, * - 10%.

We can see that the coefficient on real wage is positive in all the specifications and statistically significant in five of the nine cases. This result is the opposite of similar income-conflict regressions reported in the literature and supports the first implication derived from the model that conflict increase with wage at least when wages are low. Gomes (2015), using the same conflict data but over a longer time period, 1979-2009, find a negative correlation between conflict and per capita consumption. The important difference is that per capita consumption is a much more aggregate measure that includes all sections of the population in the district, while agricultural wages only capture the wages for agricultural labourers, the poorest class of workers. Also, here we are estimating short term effects as compared to Gomes (2015) who is looking for longer term effects.

We will now briefly discuss the coefficient estimates of the other regressors. We have already commented on the lagged dependent variable and how it proxies for the district level fixed effect. The only control variable available as a panel is forest cover and its coefficient estimates are mostly positive but significant only for pooled Poisson and Arellano-Bond GMM regressions. Forest cover can be understood to affect the conflict in various ways. More forest cover help guerrilla warfare and tilts the scales in favour of the Maoists, while higher forest cover may also mean that the productivity of natural resources for traditional use is higher. It is not very clear if these effects would lead to increase or decrease in conflict. Maoist presence in 2001-05 has a consistently positive and significant effect on conflict as would be expected. The effect of mining is negative and significant in three specifications. This appears counter-intuitive but it is to be noted that the variable only denotes the presence of mining and not the start of new mines, which might be expected to increase conflict. The existence of old mines might just mean that there is no scope for starting more mines, hence no incentive for the state to acquire forest land, and hence no reason for conflict thus explaining the negative coefficient. The Scheduled Tribe population share has a positive and significant coefficient estimate in three of the four specifications. This again may appear counter-intuitive as increased population share should mean increased political power of the tribal population and hence less conflict. But the tribal population is not randomly distributed. The population share of tribals is likely to be higher in areas that have more natural resources to start with and hence more potential for conflict. The last control variable Urbanisation is not found to be significant in any specification.

5.1 Addressing endogeneity

An important source of bias that we have not discussed yet is the bias because of reverse causality. The occurrence of conflict will affect production and hence wages, making wages endogenous to conflict. But even in the presence of such endogeneity, one would

Table 3: First stage regression showing NREGA implementation as a relevant instrument

Dependent variable is *Real wage$_{it}$*	
	1
	FE linear
NREGA$_{it-2}$	6.268 ***
	(1.002)
Forest cover$_{it}$	- 0.201
	(0.100)
Number of observations	524
Number of districts	130
F-statistic for testing coefficient of NREGA$_{it-2}$=0	39.11
p-value of F-statistic	0.000

Standard errors are clustered at the district level and are given in parentheses.
Significance levels: *** - 1%. ** - 5%, * - 10%.

expect the bias to be negative, i.e. conflict to reduce wages, and hence the coefficients here may be underestimating the unbiased effect of wages on conflict. If the direction of bias is negative then it does not undermine our result about rising wages leading to increase in conflict.

The other source of bias is measurement error. As described earlier, the wage data is noisy and this noise can lead to attenuation bias. This would also imply a negative bias and would make the coefficients in Table **??** lower bounds for the actual coefficient.

We address the endogeneity as well as the measurement error by instrumenting for real wage. We require an instrument that is relevant, i.e. significantly correlated with real wage, and exogenous, i.e. not influencing conflict in any way except through wages. The implementation of the National Rural Employment Guarantee Act (NREGA) is a potential instrument.

This employment programme was implemented from 2006-08 across rural India providing 100 days of paid work at a wage that was generally higher than the prevailing rural wage. The scheme was rolled out across the districts of India in three phases, which provides the required variation for using it as a variable. A number of studies (Berg et al., 2013; Azam, 2012; Imbert and Papp, 2015) have found that NREGA implementation did raise wages in rural India. Using data from the Ministry of Rural Development, we construct a binary variable for NREGA implementation which has value 1 for a district in the year the implementation was started and for all years thereafter. The correlation between NREGA implementation and real wage was highest for a 2 year delay. This could be due to the time taken for the programme to be implemented and general equilibrium effects to occur to have an effect on the wage levels. This also accounts for the fact that the collection period for the annual wage data for year t starts in July of year t-1. The selection of districts for the different phases of the scheme was not random and hence NREGA implementation cannot serve as an instrument for cross-sectional data. But if we remove district fixed effects, then the district-specific time-invariant factors that led to the selection of the district in a particular phase are eliminated, thus making NREGA implementation exogenous to conflict and hence a valid instrument for fixed effect estimation. Also, NREGA is funded by the central government and hence its implementation should not affect the budgets of the state governments who are the 'state' in the context of the model, which decides on deploying security forces. The first stage linear regression showing the relevance of the instrument is given in Table **??**. It is evident that NREGA implementation is highly correlated with real wage.

The exclusion restriction can be questioned as NREGA may impact conflict through other channels such as information cost (Berman et al., 2011; Khanna and Zimmermann, 2014). We find that NREGA is significantly correlated with conflict only after two years, when it has had a significant effect on wages. Hence, it is unlikely that NREGA would directly effect conflict through a channel other than wages. Table 4 presents results of reduced form regressions of number of incidents on NREGA. Current, lagged and twice lagged values of *Real wage* and *Number of incidents* are regressed on twice lagged NREGA and current real wage and number of incidents are regressed on current and lagged values of NREGA. We find that only significant correlation is between current values of number of incidents and twice lagged NREGA in the log-linear specification. It also shows that NREGA implementation affects the number of incidents only when it has a substantial effect on real wage, thus supporting the exclusion restriction.

Table 5 presents the results of regressions using NREGA implementation as an instrument for *Real wage*. The relationship between *NREGA* and *Real wage* can be assumed to be linear without any problems since real wage is a continuous variable. But to use this linear relationship to instrument for *Real wage* in a non-linear regression is not straight forward. As a first approximation, we assume the relationship between *Real wage* and *Number of incidents* to be linear and conduct a simple two stage least squares with fixed effects (2SLS FE). The results of this are presented in the first column. The coefficient estimate for *Real wage* is positive but not significant.

A simple way to account for the exponential relationship between the dependent and the explanatory variables as a result of the Poisson distribution would be to use the logarithm of *Number of incidents* as the dependent variable and conduct a log-linear regression (Hausman et al., 1984; Blundell et al., 1999). The only problem with this approach is the presence of a large number of zeroes. This can be accounted for in the following way. Re-

Table 4: Reduced form regression of number of incidents with NREGA implementation

Panel A					
Dependent variable	Real wage in district i and year				
	t	t-1	t-2	t	t
	1	2	3	4	5
	FE linear	FE linear	FE linear	FE linear	FE linear
$NREGA_{it-2}$	6.282 ***	1.720 *	- 0.977		
	(1.009)	(0.889)	(1.119)		
$NREGA_{it-1}$				2.962 ***	
				(0.885)	
$NREGA_{it}$					1.195
					(1.459)
Number of observations	524	435	333	524	524
Number of districts	130	130	126	130	130

Panel B					
Dependent variable	Number of incidents in district i and year				
	t	t-1	t-2	t	t
	1	2	3	4	5
	FE log-linear	FE log-linear	FE log-linear	FE log-linear	FE log-linear
$NREGA_{it-2}$	0.121 ***	0.021	- 0.037		
	(0.038)	(0.035)	(0.063)		
$NREGA_{it-1}$				0.019	
				(0.034)	
$NREGA_{it}$					- 0.042
					(0.052)
Number of observations	740	592	444	740	740
Number of districts	148	148	148	148	148

The coefficients in Panel A are marginal effects while those in Panel B are semi-elasticities. Standard errors are clustered at the district level and are given in parentheses. Significance levels: *** - 1%. ** - 5%, * - 10%.

Table 5: Regressions using NREGA implementation as an instrument for*Real wage*

Dependent variable is *Number of incidents*$_{it}$			
	1	2	3
	FE 2SLS linear	FE 2SLS log-linear	GMM Wooldridge
Coefficients	ME	SE	SE
Real wage$_{it}$ (Rs/day)	0.020	0.020 ***	0.057 *
	(0.085)	(0.008)	(0.030)
Forest cover$_{it}$	- 0.055***	- 0.011	0.069 *
	(0.148)	(0.027)	(0.029)
Zero$_{it}$		- 0.553***	
		(0.065)	
Number of observations	518	518	393
Number of districts	124	124	123

ME denotes marginal effect and SE is semi-elasticity. In columns 1 and 2, districts that have only one observation are omitted. In column 3 we lose more observations because of the differencing implied in the moment condition. Standard errors are clustered at the district level and are given in parentheses. Significance levels: *** - 1%. ** - 5%, * - 10%.

place all the zeroes with ones and introduce a dummy variable *zero,* which takes a value of one whenever the number of incidents is zero. This specification can now be estimated with 2SLS FE. The results are presented in the second column of Table 5. The coefficient estimate for *Real wage* is positive and significant and and its magnitude is larger than that in the earlier Poisson regressions, supporting the argument that the bias caused by reverse causality is negative.

The procedure of replacing zeroes with ones is clearly not satisfactory. The problem is to account for fixed effects while addressing endogeneity through a linear instrument applied to a Poisson specification. This problem has been addressed by Wooldridge (1997), who provides a transformation of the specification that yields a moment condition which eliminates the fixed effects.[9] Windmeijer (2000) suggested that we express the explanatory variables as deviations from their overall mean to allow specifications with non-negative explanatory variables to be estimated using GMM. We use the Wooldridge transformation with the Windmeijer correction to estimate our specification. The results are presented in the third column. The coefficient estimate for real wage is positive and significant at 10%. Note that we lose a large number of observations because of the dif-

[9]If the specification is $y_{it} = exp(x_{it}\beta + \delta_i + \varepsilon_{it})$, with z_{it} being the vector of instruments for x_{it} , then the moment condition is $E\left(\frac{y_{it}}{exp(x_{it}\beta)} - \frac{y_{it-1}}{exp(x_{it-1}\beta)} \mid z_{it}\right) = 0$.

ferencing that is implied by the moment condition. The magnitude of the estimate at 0.057 is again larger than previous Poisson estimates (a Poisson regression on the same subsample as the Wooldridge regression yielded a coefficient of 0.029, significant at 1%) and it implies that an increase in the real wage by one rupee would increase the number of incidents by 5.7%. This means that one standard deviation change in the real wage, which is Rs 14.5, can lead to an increase in the number of incidents by more than 80%. With the mean number of incidents being 2.54, this is equivalent to causing an additional two incidents. Another way of looking at the impact is through the effect of NREGA. Table 3 shows that two years after the start of NREGA implementation, real wages increased on an average by Rs 6.3. This translates to an increase of 0.9 incidents. So this can be interpreted as saying that at the average number of incidents, the increase in wages because of the implementation of NREGA lead to almost 1 additional incident. If we compare these estimates to those in Table 2, we an see that the bias was negative as expected.

5.2 Non-linearity

So far we have tested and found that the data shows a positive effect of real wages on conflict. The proposition derived from the model states that the effect is positive when wages are low. We have argued that the wages are indeed low as they are around the legal minimum wage, but it would still be instructive to test if the coefficient denoting the effect of real wage on conflict changes with the wage level. We would presume that as wages increase, the opportunity cost effect would become stronger and the coefficient would change from positive to negative. We tested this in two ways. One was by including a quadratic component of wage in the regression and the other was to include a dummy variable for wages below the median. In both cases the signs of the coefficients on the additional regressors were mostly in line with the prediction, i.e. negative for the quadratic term and positive for the low wage dummy, but the magnitudes of the coefficients were not statistically significant. The most probable reason is that the range of wages available in the sample is not enough to extract the changes in the coefficient, and a larger sample with either more observations or with a larger range would be needed to conclusively test this aspect of the proposition. The results of the regressions are not shown here and are available on request.

5.3 Alternative explanations

It is possible to think of alternative explanations for the results shown above. One is that the conflict is causing an increase in wages. Increased recruitment by the Maoist may lead to increase in the demand for labour and hence increase wages. The IV regression does not support this hypothesis and it is highly unlikely that recruitment by the Maoists, (which might be in hundreds in any year given that their total number is a few thousands) can affect the wage rates significantly. The other explanation could be that the Maoists are deliberately targeting areas with higher wages for reasons other than those described in the model, possibly as they could extort more money. These wages are for manual agricultural work and are barely enough for subsistence, hence it is unlikely that the Maoists will be extorting money from these people. Another explanation could be that as the dependent variable is based on news reports, it may be the case that incidents from districts with higher wages get reported more frequently. This selection bias is possible, but the selection is more likely to be dependent on urbanisation of the district rather than the wage level and we find no effect of urbanisation on the dependent variable. Hence, this explanation is not plausible.

5.4 Robustness checks

We perform a number of robustness checks on the results. As mentioned earlier, the wage data has a lot of missing values. To check against possible bias due to sytematic non-reporting, we conduct regressions on a subsample only including 96 districts that consistently report wages and find that the results don't change. We also conudct regressions excluding the state of Bihar, where the conflict is more along caste lines, and excluding the year 2010, the year for which conflict data was collected by the author, respectively and find that the results are robust. Finally, we check the results for a slightly different definition of wage. We only take the real wages for the unskilled tasks like herding and field labour as the explanatory variable. We find that the coefficient estimates are similar to previous regression although the significance levels are a little lower probably because of greater measurement error.

Since the instrument variable is a binary variable that is 0 before the implementation of the program and 1 after it, it can be argued that the results just reflect different trends in the districts that were chosen for the three phases. On including phase-specific linear trends we find that the instrument is still relevant with an F-statistic of 12. The coefficient of the log-linear IV regression remains positive and significant at 5%. The coefficient of the Wooldridge GMM regession is still positive but loses significance. This could be because of insufficient power due to smaller number of observations. The results of Khanna and Zimmerman (2014), who use a regression discontinuity design to find that NREGA implementation led to an increase in the Maoist conflict, also serve as a robustness check for the results of this paper.

The results for the robustness check regressions are not shown here and are available on request.

6 Conclusion

This paper describes a game-theoretic model of conflict over a natural resource that shows how a rational utility-maximising local population is more likely to choose conflict because of reduced risk aversion as its wages increase. The model also shows that the relationship between conflict and state capacity can be non-monotonic as the willingness of the state to fight increases with state capacity and that of the rebels decreases. The model allows for a more complex understanding of the relationship between wages and conflict, and state capacity and conflict, as it adds the mechanism of risk aversion to mechanisms already explored in the literature, like opportunity cost.

This paper is also the first to document a positive correlation between wages and conflict in central India, which contradicts the prevailing view of the relationship between wages and conflict. The advantage of this analysis over the more prevalent empirical studies of the economic causes of conflict is that it uses wage instead of income per capita and that it is a within-country study, negating some of the problems of country specific heterogeneity. However, as the analysis only covers a period of five years, the results presented here are short-term effects and the longer term effect of wage on conflict may be different.

The ultimate aim of the paper is to provide a framework for thinking about government policy. In this model, a repressive policy like Operation Greenhunt can be expected to increase conflict in some parts and reduce it in others, but it does not resolve the basic cause of the conflict. A development based policy can be expected in the long run to increase the welfare of the tribals to a point where they do not value the natural resource, but in the short run may not prove as effective. A policy like the National Rural Employment Guarantee Act which has actually increased market wages (Berg et al, 2012; Azam, 2012; Imbert and Papp, 2015) may in fact lead to increased conflict. The Forest Rights Act (FRA) seems to be the policy, under this model, that addresses the root cause of the conflict - lack of user rights on natural resources. By providing rights to tribals, the state may even enable trading between the tribals and non-tribals for these rights leading to a Pareto-superior outcome compared to the allocation by the state. More importantly, it may remove the basic cause of the conflict, after which repressive actions like Operation Greenhunt may prove much more successful in addressing the problem. The results presented here do not mean that welfare and social security programmes should not be carried out in conflict areas. The point being made is that in certain situations, as in the case of the Maoist conflict, they are not the solution to the problem and may in fact exacerbate the conflict.

References

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Azam, M. (2012). The impact of indian job guarantee scheme on labor market outcomes: Evidence from a natural experiment. IZA Discussion Paper 6548.

Bates, R. H. (2008). *When things fell apart: State failure in late-century Africa*. Cambridge University Press.

Berg, E., Bhattacharyya, S., Durgam, R., and Ramachandra, M. (2013). Can public works increase wages? Evidence from India. Working paper.

Berman, E., Callen, M., Felter, J. H., and Shapiro, J. N. (2011). Do working men rebel? insurgency and unemployment in Afghanistan, Iraq, and the Philippines. *Journal of Conflict Resolution*, 55(4):496–528.

Besley, T. and Persson, T. (2010). State capacity, conflict, and development. *Econometrica*, 78(1):1–34.

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

Blundell, R., Griffith, R., and Van Reenen, J. (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies*, 66(3):529–554.

Buhaug, H. (2006). Relative capability and rebel objective in civil war. *Journal of Peace Research*, 43(6):691–708.

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

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

Do, Q.-T. and Iyer, L. (2010). Geography, poverty and conflict in Nepal. *Journal of Peace Research*, 47(6):735–748.

Eynde, O. V. (2011). Targets of violence: Evidence from Indiaï¿œs Naxalite conflict. Job market paper, LSE.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(01):75–90.

Fetzer, T. (2014). Social insurance and conflict: Evidence from India. Job market paper, LSE.

Fjelde, H. and De Soysa, I. (2009). Coercion, co-optation, or cooperation? State capacity and the risk of civil war, 1961 2004*. *Conflict Management and Peace Science*, 26(1):5–25.

Gomes, J. F. (2015). The political economy of the Maoist conflict in India: an empirical analysis. *World Development*, 68:96–123.

Government of India (2011). Naxal hit areas. Available at http://www.satp.org/satporgtp/countries/india/maoist/documents/papers/DATA_%20Extent_Naxal_Violence.pdf. Accessed: 28-01-2016.

Grossman, H. I. (1991). A general equilibrium model of insurrections. *The American Economic Review*, 81(4):912–921.

Guha, R. (2007). Adivasis, Naxalites and Indian democracy. *Economic and Political Weekly*, pages 3305–3312.

Hausman, J., Hall, B. H., Griliches, Z., et al. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, 52(4):909–38.

Hegre, H. and Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of conflict resolution*, 50(4):508–535.

Hendrix, C. S. (2010). Measuring state capacity: Theoretical and empirical implications for the study of civil conflict. *Journal of Peace Research*, 47(3):273–285.

Himanshu (2005). Wages in rural India: Sources, trends and comparability. *Indian Journal of Labour Economics*, 48(2).

Hirshleifer, J. (1988). The analytics of continuing conflict. *Synthese*, 76(2):201–233.

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

Imbert, C. and Papp, J. (2015). Labor market effects of social programs: Evidence from India's employment guarantee. *American Economic Journal: Applied Economics*, 7(2):233–263.

Indiastat (2015). http://www.indiastat.com.

Kapur, D., Gawande, K., and Satyanath, S. (2012). Renewable resource shocks and conflict in India's Maoist belt. Center for Global Development Working Paper.

Khanna, G. and Zimmermann, L. (2014). Fighting Maoist violence with promises: Evidence from India's employment guarantee scheme. *Economics of Peace and Security Journal*, 9(1):30–36.

Labour Bureau, Government of India (2013). Statistical data for labour. Available at http://labourbureau.nic.in/indtab.pdf. Accessed: 28-01-2016.

Miguel, E., Satyanath, S., and Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, 112(4):725–753.

Ministry of Agriculture, Government of India (2006-09). Agricultural wages of India.

Ministry of Environment and Forests, Government of India (2005-11). Forest survey of India.

Ministry of Law and Justice, Government of India (2006). The Scheduled Tribes and Other Traditional Forest Dwellers (Recognition of Forest Rights) Act.

Özler, B., Datt, G., and Ravallion, M. (1996). A database on poverty and growth in India. Technical report, The World Bank.

Padel, F. and Das, S. (2006). Anthropology of a genocide: Tribal movements in central India against over-industrialisation. *Report for the SAAG*.

Planning Commission of India (2008). Development challenges in extremist affected areas.

Ramesh, J. (2012). 21st century Maoism in India. Nehru Memorial Museum and Library Public Lecture. Available at https://www.youtube.com/watch?v=q03ghjsebIk26m45s.

Reserve Bank of India (2008-12). State finances: A study of budgets.

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system gmm in stata. *Center for Global Development working paper*, (103).

SATP (2015). South Asian Terrorism Portal. http://www.satp.org.

Shrivastava, A. and Kothari, A. (2012). *Churning the earth: The making of global India*. Penguin.

Simeon, D. (2010). Permanent spring. *Seminar*, 607, March.

Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *American Economic Review*, 82(4):720–739.

Windmeijer, F. (2000). Moment conditions for fixed effects count data models with endogenous regressors. *Economics Letters*, 68(1):21–24.

Windmeijer, F. (2006). Gmm for panel count data models. *Cemmap Working Paper No. CWP21/06*.

Wooldridge, J. M. (1997). Multiplicative panel data models without the strict exogeneity assumption. *Econometric Theory*, 13(05):667–678.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics*, 90(1):77–97.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

A Appendix I

In order to simplify notation, we follow the convention that X^Y represents $\dfrac{\partial X}{\partial Y}$ and X^{YZ} represents $\dfrac{\partial^2 X}{\partial Y \partial Z}$. We also denote $g(c(w,\beta,s))$ as $g(w,\beta,s)$.

Proof of Lemma 1: The derivative of P_c with respect to w is given by

$$P_c^w = \frac{(g_T^w(w,0,0) - g_T^w(w,0,1))(g_T(w,1,1) - g_T(w,0,1)) - (g_T(w,0,0) - g_T(w,0,1))(g_T^w(w,1,1) - g_T^w(w,0,1))}{(g_T(w,1,1) - g_T(w,0,1))^2}$$

Since we are concerned with the sign of P_c^w, we can look only at the numerator. For $P_c^w < 0$, we need

$$\begin{aligned} &(g_T^w(w,0,0) - g_T^w(w,0,1))(g_T(w,1,1) - g_T(w,0,1)) \\ &\qquad - (g_T(w,0,0) - g_T(w,0,1))(g_T^w(w,1,1) - g_T^w(w,0,1)) < 0 \\ \mathit{Or},\quad & g_T^w(w,0,0)[g_T(w,1,1) - g_T(w,0,1)] - g_T^w(w,1,1)[g_T(w,0,0) - g_T(w,0,1)] \\ &\qquad - g_T^w(w,0,1)[g_T(w,1,1) - g_T(w,0,0)] < 0 \end{aligned}$$

Let us first show this for the assumed functional form. We can substitute the expressions for $g_T(.)$ and $g_T^w(.)$.

$$\frac{1}{\gamma(w)^{1/\gamma}}[a]-\frac{1-h_1}{\gamma(a+(1-h_1)w-h_2)^{1/\gamma}}[wh_1-h_2]-\frac{1-h_1}{\gamma((1-h_1)w-h_2)^{1/\gamma}}[a-wh_1-h_2]<0$$

Since $h_1<1$ and $w<(a-h_2)/h_1$, we can rewrite this as

$$\left[\frac{1}{\gamma(w)^{1/\gamma}}[a]-\frac{1-h_1}{\gamma(a+(1-h_1)w-h_2)^{1/\gamma}}[wh_1-h_2]\right]\frac{1}{(1-h_1)(a-wh_1-h_2)} < \frac{1}{\gamma((1-h_1)w-h_2)^{1/\gamma}}$$

Now, as $w\to w_c=\dfrac{h_2}{1-h_1}$, the right hand side will tend to infinity, while all terms on the left hand side are finite. Hence at some $\tilde{w}>w_c$, this inequality will be satisfied.
We now prove the lemma for the general case. We need

$$g_T^w(w,0,0)[g_T(w,1,1)-g_T(w,0,1)]-g_T^w(w,1,1)[g_T(w,0,0)-g_T(w,0,1)] -g_T^w(w,0,1)[g_T(w,1,1)-g_T(w,0,0)]<0$$

Using A6 $[c(w,1,1)>c(w,0,0)]$ and A1 $[g'(c)>0]$, we can rewrite this as

$$\frac{g_T^w(w,0,0)[g_T(w,1,1)-g_T(w,0,1)]-g_T^w(w,1,1)[g_T(w,0,0)-g_T(w,0,1)]}{[g_T(w,1,1)-g_T(w,0,0)]}<g_T^w(w,0,1)$$

Now,

$$g_T^w(w,0,1)=\frac{\partial g}{\partial c}\frac{\partial c}{\partial w}\mid_{c=c(w,0,1)}$$

Using A3, we know that $\dfrac{\partial c}{\partial w}\mid_{c=c(w,0,1)}>0$. Hence, we can rewrite the required inequality as

$$\frac{g_T^w(w,0,0)[g_T(w,1,1)-g_T(w,0,1)]-g_T^w(w,1,1)[g_T(w,0,0)-g_T(w,0,1)]}{[g_T(w,1,1)-g_T(w,0,0)]}\frac{1}{\frac{\partial c}{\partial w}\mid_{c=c(w,0,1)}} < \frac{\partial g}{\partial c}\mid_{c=c(w,0,1)}$$

Using A4, as $w\to w_c$, $c(w,0,1)\to 0$ and using A1, as $c(w,0,1)\to 0$, $\dfrac{\partial g}{\partial c}\mid_{c=c(w,0,1)}\to\infty$. From A1, A2, A5 and A6, we can see that the left hand side is finite. Hence there will be a $\tilde{w}>w_c$, for which the inequality is satisfied.

B Appendix II

Proof of proposition 1: We know that the interior solution to the rebel's optimisation problem is given by

$$P^R(F,\hat{R}(F))=1/v$$

Using the implicit function theorem, we get

$$\frac{\partial \hat{R}(F)}{\partial F} = -\frac{P^{RF}}{P^{RR}}$$

We now look at how $P(\hat{R}(F),F)$ changes as we change F.

$$\frac{dP(F,\hat{R}(F))}{dF} = P^F + P^R\frac{\partial \hat{R}(F)}{\partial F} = P^F + P^R\left(-\frac{P^{RF}}{P^{RR}}\right)$$

Using A9, we get

$$\frac{dP(F,\hat{R}(F))}{dF} < 0$$

Since, $\hat{R}(F)$ depends only on v, hence for given values of v, there exists an F_c such that $P(F,\hat{R}(F)) \le P_c$ for all $F \ge F_c$. We know that it is optimal for T to not support the rebels if $P \le P_c$. Hence, there exists an F_c such that it is optimal for T to choose $s = 0$ for all $F \ge F_c$.

F_c is given by

$$P(F_c,\hat{R}(F_c)) = P_c$$

Writing F_c as $F_c(P_c)$ and using implicif function theorem, we get

$$\frac{\partial F_c}{\partial P_c} = -\frac{-1}{\dfrac{dP(F,\hat{R}(F))}{dF}} < 0$$

Using Lemma 1, we get

$$\frac{\partial F_c}{\partial w} = \frac{\partial F_c}{\partial P_c}\frac{\partial P_c}{\partial w} > 0$$

C Appendix III

In stage 1, the state's problem is

$$\max_{0\le F\le e} U_M = (1 - P(F,S^*(F)R^*(F,S^*(F))))\,k + f(e-F)$$

Substituting for $S^*(F)$ and $R^*(F,S^*(F))$, we get three scenarios

$$\max_{0\le F\le e} U_M = \begin{cases} f(e) & if\ F = 0 \quad \ldots(I) \\ (1-P(F,\hat{R}(F))k + f(e-F) & if\ 0 < F < F_c \quad \ldots(II) \\ k + f(e-F) & if\ F \ge F_c \quad \ldots(III) \end{cases}$$

Note that if $F_c > E$, then the third scenario is ruled out. If $F_c \le E$ then it is evident that the optimal value of F in scenario (III) is F_c. For scenarios (I) and (II) let us define $\hat{F}$ as

$$\hat{F} = arg \max_{0\le F\le e} (1-P(F,\hat{R}(F))k + f(e-F)$$

The rebels' best response $\hat{R}(F)$ is given by

$$P^R(F,\hat{R}(F)) = 1/v$$

The first order condition of the state's objective function is

$$\frac{\partial U_M}{\partial F} = -\frac{d\,P(F,\hat{R}(F))}{dF}k - f'(e-F) = 0$$

Hence, for an interior solution to the state's problem and as a result for an equilibrium to exist, we need $\frac{d\,P(F,\hat{R}(F))}{dF} < 0$,which we have shown in Appendix II.

The sufficient condition for an interior solution is

$$-k\frac{d\,P(F,\hat{R}(F))}{dF}\,|_{F=0} - f'(e) > 0 > -k\frac{d\,P(F,\hat{R}(F))}{dF}\,|_{F=e} - f'(0)$$

$$Or, \quad \frac{f'(e)}{-\frac{d\,P(F,\hat{R}(F))}{dF}\,|_{F=0}} < k < \frac{f'(0)}{-\frac{d\,P(F,\hat{R}(F))}{dF}\,|_{F=e}}$$

Now let us look at U_M in the three scenarios.

$$\begin{aligned} U_M(I) &= f(e) \\ U_M(II) &= (1-P(\hat{F},\hat{R}(\hat{F})))k + f(e-\hat{F}) \\ U_M(III) &= k + f(e-F_c) \end{aligned}$$

The state will prefer *I* if $U_M(I) \geq U_M(II)$ and $U_M(I) \geq U_M(III)$.This implies

$$k \leq min\left\{ f(e) - f(e-F_c), \frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \right\}$$

The state will prefer *II* if $U_M(II) \geq U_M(I)$ and $U_M(II) \geq U_M(III)$.This implies

$$\frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \leq k \leq \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))}$$

The state will prefer *III* if $U_M(III) \geq U_M(I)$ and $U_M(III) \geq U_M(II)$.This implies

$$k \geq max\left\{ f(e) - f(e-F_c), \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))} \right\}$$

Hence, the state's equilibrium strategy is

$$F^* = \begin{cases} 0 & if\ k \leq min\left\{ f(e) - f(e-F_c), \frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \right\} \\ \hat{F} & if\ \frac{f(e)-f(e-\hat{F})}{1-P(\hat{F},\hat{R}(\hat{F}))} \leq k \leq \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))} \\ F_c & if\ k \geq max\left\{ f(e) - f(e-F_c), \frac{f(e-\hat{F})-f(e-F_c)}{P(\hat{F},\hat{R}(\hat{F}))} \right\} \end{cases}$$

D Appendix IV

Let us rewrite the Repression-Conflict margin as follows:

$$RC = [k + f(e - F_c)] - [(1 - P(\hat{F}, \hat{R}(\hat{F})))k + f(e - \hat{F})]$$

If $RC > 0$, we have Repression and if $RC < 0$ we have Conflict. Now we check how RC changes with$e$.

Using the envelope theorem, we know that

$$\frac{d([(1 - P(\hat{F}, \hat{R}(\hat{F})))k + f(e - \hat{F})])}{de} = f'(e - F)$$

Hence,

$$\frac{dRC}{de} = f'(e - F_c) - f'(e - \hat{F}) > 0$$

This is because $\hat{F} < F_c$ and $f(.)$ is concave. Therefore an increase in state capacity can change the equilibrium from Conflict to Repression.

We can rewrite the Peace-Conflict margin as follows:

$$PC = f(e) - [(1 - P(\hat{F}, \hat{R}(\hat{F})))k + f(e - \hat{F})]$$

If $PC > 0$, we have Peace and if $PC < 0$ we have Conflict. Now we check how PC changes with e.

$$\frac{dPC}{de} = f'(e) - f'(e - \hat{F}) < 0$$

This is because $\hat{F} > 0$ and $f(.)$ is concave. Therefore an increase in state capacity can change the equilibrium from Peace to Conflict.