2007/81

■

Does inequality make us rebel?
A renewed theoretical model applied to South Mexico

Jean-François Maystadt

CORE DISCUSSION PAPER
2007/81

Does inequality make us rebel? A renewed theoretical model applied to South Mexico

Jean-François MAYSTADT[1]

October 2007

Abstract

Since Collier and Hoeffler (1998, 2004), it has been supported that inequality, measured at national level, does not affect the risk of conflict. Based on a renewed theoretical framework, the purpose of the paper is to explore the role of inequality in localized conflicts. We argue that previous findings might be biased by the myopic nature of cross-country analysis. Consistently with the model, Probit estimations indicate that income inequality measured at municipal level was significant in motivating people to support the rebellion in South Mexico. At this geographical level, we also find an increase in income per capita could exacerbate the risk of conflict in a situation where the rebel leader would have greater incentives to loot the local production compared to the opportunity cost associated with fighting for the worker.

Keywords: rebellion, inequality, income, Mexico.

JEL Classification: O18, O54, C35

---

[1] CORE, Université catholique de Louvain, Belgium and FNRS.
E-mail: jean-francois.maystadt@uclouvain.be

I am grateful to Dave Hodson from CIMMYT (International Maize and Wheat Improvement Center) for providing geographic data and to Kristian Behrens, Bart Cockx, Giordano Mion, Dominique Peeters, Walter Steingress, Jacques Thisse, Vincenzo Verardi and Shlomo Weber as well as participants to UCL DWs (May 2006 and January 2007) and to the Winter School on Polarization and Conflict (Cordoba, November 2006) for helpful comments. I also thank Caridad Araujo, David Dabin and Claire Dujardin for their advices on the way to handle geographic data. I acknowledge financial support from the Fond National de la Recherche Scientifique (FNRS). Any remaining errors are naturally my own.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the author.

1 Introduction

It is commonly argued that increasing inequality leads to conflict or revolution. This argument has been deepened by Gurr (1970) who supports that collective violence was driven by relative deprivation, defined as the difference between what a social group believes it deserves and what it really gets to live. Even Sen (1997, 1) suggests in his exposed motivations to study further the issue of inequality that "the relation between inequality and rebellion is indeed a closed one, and its runs both ways. That a perceived sense of inequity is a common ingredient of rebellion in societies is clear enough". However, this relation seems to be nothing more than an assertion. Russett (1964) and Muller (1985) already underlined how this argument is deeply rooted in the works of authors such as Karl Marx or Alexis de Tocqueville, but clearly lacks empirical support. [1] With the means of non-linear models, Collier and Hoeffler (1998) have seminally renewed the empirical analysis on the causes of conflict. Extending their first analysis, Collier and Hoeffler (2004) found that measures of social grievances such as income inequality do not systematically affect the risk of conflict. By sharp contrast, measures of greed such as economic decline are found to be significant. [2] Collier (2000, 10) concludes that "inequality does not seem to affect the risk of conflict. Rebellion does not seem to be the rage of the poor". The non-significance of income inequality measured by the Gini coefficient is corroborated by Fearon and Latin (2003). This result has been much debated by authors such as Nafziger and Auvinen (2002) or Stewart (2000) who find a positive relationship between income inequality and the occurrence of civil war. However, the period investigated by Nafziger and Auvinen (2002) is clearly shorter than the other studies and no detailed results are provided to support their statement. Stewart (2000) adopts another type of research design as she illustrates her concept of horizontal inequality through the study of Uganda and widens the notion of inequality to social and political dimensions. Comparison is difficult as she did not seem to test econometrically the significance of her estimates of horizontal inequalities. The purpose of this paper is then to explore how inequality could motivate people to rebel or at least, to support a rebellion by studying its role in South Mexico.

The paper argues that finding no significance of income inequality is not highly surprising and is even misleading. On the one hand, despite their theoretical foundations emphasizing the cost-benefit calculus undertaken by individuals, these cross-country studies overlook individual motivations. Local factors are likely to be determinant in explaining localized or intra-national insurrections. For instance, we can doubt that the decision to support the rebellion in South Mexico will be as much influenced by the difference of welfare with someone

[1] As far as we know, Russett (1964) is the first to have approached this relationship by using the Gini coefficient as a measure of the total inequality of a distribution. Both authors use multiple regressions to emphasize that several factors might influence an uprising.

[2] Since Collier and Hoeffler (2004), distinguishing greed and grievance has become common in the economic literature on conflict. Grievance is defined as "a motivation based on a sense of injustice in the way a social group is treated, often with a strong historical dimension" while greed, an economic opportunity but characterized as "an acquisitive desire similar to crime, albeit often on a much larger scale"(Murshed, 2002, 189). This paper aims at going beyond this debate by considering how the possible greedy leaders might manipulate the grievances of others (Keen, 2001).

living far in the North compared with someone being located in the closed neighborhood. Although bearing the obvious advantage of generalization, standard cross-country studies are too myopic to explore how local perceptions of inequality might affect the probability to support a rebellion. [3] On the other hand, most econometric analysis on the subject refer to some standard models in the economics of conflict to motivate the choice of their explanatory variables. However, empirical results are hardly shown to be consistent with the theoretical framework underlying their work. A noticeable exception is the recent paper of Olsson (2007). Furthermore, these models seem more relevant in explaining state-wide civil war. In section 2, we propose a renewed theoretical model, more adequate to study the causes of localized conflicts. We emphasize the need to approach both relative and absolute deprivations in a common framework. It also provides a possible reconciliation between an economic approach emphasizing the opportunity cost of joining a rebellion and a view more often encountered in political sciences for which larger resources increase the potential pay-offs for the warring parties. the empirical implication is that ignoring one dimension could bias the estimate of the other. Also dealing with other sources of endogeneity, section 3 empirically tests how income inequality and income per capita could motivate people to rebel or at least, to support the rebellion by studying its role in South Mexico. Our results are consistent with the main predictions of the theoretical model.

To check the robustness of our analysis, other factors of conflict would need to be tested. Of particular interest in the literature is the role of ethnic distinction in motivating people to fight. Although the primordialist approach predicting a struggle among old identities following the end of the Cold War has been dismissed due to its simplistic conception of ethnicity, the formation of groups often follows ethnic distinctions. Perceived or real differences between groups are often used or manipulated by leaders for mobilization. A relative consensus in the literature seems to emerge according to which ethnic distinctions alone are not sufficient to bring groups to violent mobilization but can be used as a resource to gain political power and economic resources (Turton, 1997). These distinctions remain crucial when combined with sources of political, social or economic grievances, even when they are socially and historically constructed or exaggerated. Section 3.4.2 will discuss the robustness of our results to alternative measurements of ethnicity and to the inclusion of other control variables.

2 The model

Most empirical works on the subject usually make some distant reference to the standard models in the economics of conflict such as Grossman (1991, 1999), Hirshleifer (1995, 2001) or Skaperdas (1992). However, consistency of the empirical results with theoretical foundations are rarely verified. For example, based on Grossman (1991, 1999), Collier and Hoeffler (2004) found that income per capita, assumed to be a proxy for greed, explains better the risk of conflict than income inequality which is assumed to be a proxy for grievance. However,

[3] Cramer (2003) also argues that national measures of inequality can also be misleading, particularly in the case of low-intensity guerilla insurrection or localized rebellion. However, contrarily to this author, we do not reject the quantitative approach to inequality measurement.

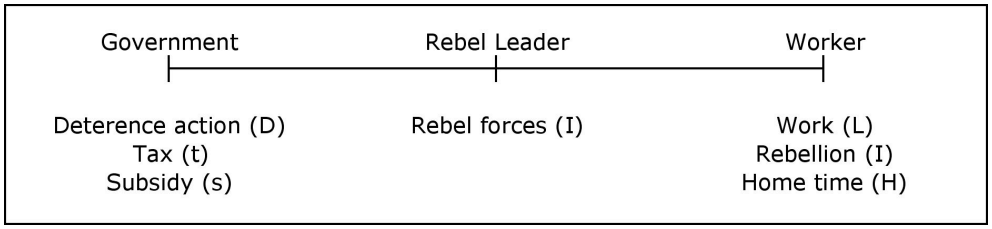
the same Grossman (1999) showed theoretically that the probability of conflict would only depend on the realization of a random variable reflecting the relative effectiveness of the rebels compared to the army soldiers. Consistency with the theory would have implied that income per capita should be insignificant, too. [4] Furthermore, the superiority of the greed variable on the grievance variable results from the comparison of two models, as the income per capita and income inequality are estimated in distinct regressions. Given the fact that one variable could impact on the other, it raises some obvious problems of potential endogeneity. At least, we need a theoretical model that would explicitly distinguish the effects of absolute wealth and inequality.

The model is inspired by Grossman (1999) and has a similar general equilibrium flavor. However, three major changes seek to make it more applicable to localized conflicts. First, the army recruitment is not done within the conflicting area. Being a soldier is not one of the choices of the worker. On the contrary, Grossman (1999) allows the worker to be a soldier and a rebel at the same time. Since the recruitment is then done uniquely within the conflicting area, the cost or return of recruiting an additional rebel is then always equal to the cost or return of recruiting an additional soldier. This simplifies the computation but also cancels out some interesting effects. Departing from this assumption will allow the level of wealth to have a differentiated effect on the level of deterrence and on the potential rebel forces. A second major difference refers to the behavior of the rebel leader. The leader is recognized as an entrepreneur but, contrary to Grossman (1999), he is not benevolent in the sense he does not necessarily maximize the welfare of his supporters. This view is at the core of many political economic analysis of conflict where the economic agenda of the rebel leader has often overcome political motivates. War is, above all, an instrument to achieve other means but is not an end in itself. Many rebel leaders have accumulated a massive amount of resources in wartimes. An obvious candidate for looting is the appropriation of natural resources such as diamonds in e.g. Angola, Congo and Liberia (Fairhead, 2000; Billon, 2001; Olsson, 2007) or Sierra Leone (Richards, 1996). Other economic motives could be observed such as the control of illegal trade, e.g. drug smuggling in Afghanistan (Rubin, 2000), the control of land in Somalia and Iraq (Keen, 1993) or the exploitation of cheaper labour in Sudan (Keen, 1994). The modeling implication is that the rebel leader will simply maximize the rent he can capture in wartimes. Similar to Grossman (1999) and contrary to Roemer (1988), the rebel leader is able to exclude non-participants from benefiting from its fighting loot but his objective is to maximize its expected net income, not necessarily the one of his supporters. Another implication is that the rebel leader does not need a final victory against the ruler to receive a positive pay-off from warfare. This conforms with the general observation that parties engaged in conflict may benefit from persistent warfare. Contrary to the conventional wisdom of two parties fighting to win against each other, war has paradoxically lead to some kind of cooperation between the parties. War could persist under equilibrium, the two fighting parties actually colluding at the expense of the most

[4] The interpretation given by Fearon and Latin (2003) of the income per capita as a proxy for the state's overall financial, administrative, police and military capabilities would be more consistent with the underlying theoretical model. Before all, it underlines the need to stick to a model at the time of interpreting the coefficient of such a variable.

vulnerable citizens. We will be mostly interested in these situations where no side is able to completely eradicate the other one. Finally, we introduce distributional considerations, by allowing the government to give a financial incentive to the potential fighter to work rather than fight. Such possibility to pay for peace is consistent with the theoretical models of Azam (1995), Azam and Mesnard (2003) and Noh (2002). According to these authors, the government does not only use the military forces to deter conflict but could actually combine it with a distributive policy in favor of their opponents. The first authors report the example of Houphouët-Boigny, former president of Côte-d'Ivoire who taxes his own ethnic group, the Akan cocoa and coffee growers to finance public infrastructure and other redistributive policies in favor of other ethnic groups. Despite some common feature, our model differs from Noh (2002) in the way we model the transfer from the ruler to the worker and not directly to the rebel leader. Therefore, we restrict the peace-enhancing role of redistribution, as fostering local production also increases the potential loot of the rebel leader.

Figure 1: Sequence of decisions



The model identifies three decision-making agents : a representative worker, a unique rebel leader and a central government. The model assigns particular objectives to these agents, the strategies available to them and their constraints. As shown in figure 1, these agents make their decisions sequentially. In this game, the government is the Stackelberg leader. It will first decide the level of deterrence, the level of tax and the labour subsidy in order to maximize the total wealth produced on its territory. Then, the rebel leader will decide to recruit or not some rebels. Finally, the worker will decide how to allocate his time. At the end of the sequence, the local producer simply adjusts, on a mechanistic way, its production and the wage according to the supply of labour.

2.1 The worker-fighter decision

The representative worker seeks to maximize his net income. He allocates his endowment of time, normalized to 1, between working for the producers, supporting the rebellion, and working informally for himself. The support to the rebellion does not necessarily need to be fighting but could be spending time participating to meetings and protestations, spreading the propaganda of the rebellion, etc. Throughout the model, we assume a technology of production of the one-factor form : $Y = \frac{1}{\alpha} AL^{\alpha}$, with $0 < \alpha < 1$ indicating a decreasing marginal utility of work. For comparability with the model of Noh (2002), the parameter

A reflects the productivity of the production. We also assume that the local producer will only adjust the production and the wage to the time allocated to labour by the worker. Therefore, the labour wage is given by the marginal productivity of labour $AL^{\alpha-1}$. Based on these assumptions, the net income of the representative worker is represented as follows:

$$\begin{aligned} E &= (1+s)w_l L + w_i I + \gamma H \\ &s.t. \\ &\quad L + I + H = 1 \\ &\quad L, I, H \geq 0 \\ &\quad \gamma > 0 \\ &\quad 0 < \alpha < 1 \end{aligned} \tag{1}$$

L is the time allocated to working ; I is the time allocated to supporting the rebellion ; H is the time allocated to informal activities ; w_l is the labour wage ; w_i is the reward provided for supporting the rebellion ; s gives the financial incentive (subsidy) given by the government to go to work ; α is the elasticity of the marginal product of time allocated to production work ; γ is the marginal return to time allocated to informal activities. The worker will maximize its expected income with respect to L, I and H. He will take the compensation wage offered by the rebel leader, and the employment subsidy introduced by the government as given. Using the Kuhn-Tucker conditions, we obtain the following first-order conditions:

$$\begin{aligned} L &= A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}(1+s)\right]^{\frac{1}{1-\alpha}} \quad \text{or} \quad L = 0 \\ w_i &\geq \gamma \quad \text{or} \quad I = 0 \\ H &= 1 - L - I \end{aligned} \tag{2}$$

The solution gives the equilibrium value of L in reaction to s. We assume the non-negativity constraint for H is not binding. In other words, γ is large enough for this constraint not to bind in equilibrium. When the rebellion reward is larger or equal to γ, the marginal product of time allocated to informal activities, the supply of rebel forces will be perfectly elastic. Given the reaction function of the worker, we derive the labour wage and the corresponding production level.

$$\begin{aligned} w_l &= \left[\frac{\alpha}{\gamma}(1+s)\right]^{-1} \\ Y &= \frac{1}{\alpha} A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}(1+s)\right]^{\frac{\alpha}{1-\alpha}} \end{aligned} \tag{3}$$

2.2 The rebel leader decision

We assume there is only one rebel leader who has the necessary skills to motivate people to rebel. However, these skills will not be enough to overcome collective action problems. Contrary to Grossman (1999), he has to face the costs in organizing the rebellion. These coordination costs increase with the number of fighters (protesters). Furthermore, the leader is not benevolent regarding his supporters. He only maximizes his own net income under a cost constraint. The rebel leader will capture a rent even when he has not overthrown the central ruler. We follow the technology of revolution given by Grossman (1999). In this setting, $(1-\lambda)$ captures the way the rebel leader is able to loot the local production. His potential loot will be proportional to the support he gets from workers and to the efficiency of his looting technology (material and immaterial). The leader is constraint as he has to pay a compensation wage to the workers sufficiently high to make them allocate parts of their time to the support of the rebellion. The fact he pays a compensation wage rather than sharing his rent among fighters is an important difference compared to previous models. The problem of the rebel leader is as follows:

$$\begin{aligned} J &= (1-\lambda)Y - w_i I - cI \\ &\text{where} \quad \lambda = \frac{1}{1+\frac{\theta I}{D}} \end{aligned} \tag{4}$$

Y represents the output produced locally ; c covers the coordination costs faced by the rebel leader ; $(1-\lambda)$ is the potential ability of the rebel leader to capture a rent, i.e. part of Y ; θ gives the efficiency of the rebels relative to the deterrence forces. θ reflects factors such as the skills of the rebels to coordinate the rebellion, the support from a foreign government or the military advantage exogenously given by the nature of the battlefield. The rebel leader maximizes his net income with respect to I, the number of fighters (supporters) he will have to compensate. He will consider the compensation wage, the variable cost of coordinating the uprising, the technology of looting and the value of θ as given. Using the reaction functions (2) and (3), the following solution is obtained :

$$\begin{aligned} I(s,D) &= \sqrt{\frac{DA^{\frac{1}{1-\alpha}}}{\alpha\theta\left(c+\gamma\right)}\left[\frac{\alpha}{\gamma}\left(1+s\right)\right]^{\frac{\alpha}{1-\alpha}}} - \frac{D}{\theta} \quad \text{for} \quad D < \hat{D} \\ I(s,D) &= 0 \quad \text{for} \quad D \geq \hat{D} \\ \hat{D} &= \frac{\theta Y}{c+\gamma} \end{aligned} \tag{5}$$

The interior solution is found when $D < \hat{D}$, i.e. the minimum level of military action necessary to deter the potential revolutionary leader to coordinate an uprising. It is an increasing function of θ and Y in the sense the more efficient the insurgents are and the greater the potential loot is, the greater the deterrence action should be to achieve the necessary condition to deter the uprising. The inverse relationship is found with c. The

greater the coordination costs are, the lower the necessary deterrence should be to avoid war.

2.3 The government decision

The objective of the government is to maximize the total wealth produced on its territory. A specificity of this model is that the government can achieve this objective by paradoxically taxing the local producers. Indeed, they will use this tax revenue to pay for peace, that is to say, to increase the incentives of the workers to work rather than support the rebellion. The government will also decide which level of deterrence it will release at a variable cost ϕ. [5] The ruling government takes the compensation wage, the technology of looting λ and the exogenous variable θ as given. The government also considers exogenously the technology of production of the local producers. The objective function of the government is defined as follows :

$$\begin{aligned} Q &= (1-\tau)\lambda Y \\ &\text{s.t.} \quad \tau\lambda Y \geq s w_l L + \phi D \end{aligned} \tag{6}$$

The first decision made by the central government is the determination of the deterrence action, D. The government seeking to maximize the local production will react proportionally to the resources of the local producers under risk of looting. Taking into account the direct effect of D on λ and the indirect effect of D on λ via I, the equilibrium level of deterrence action, when it is lower than $\hat{D}$, is as follows :

$$D = \frac{(\gamma+c)}{4\alpha\theta\phi^2} A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}(1+s)\right]^{\frac{\alpha}{1-\alpha}} \quad \text{if} \quad 0 < D < \hat{D} \tag{7}$$

Given (5) and (7), the level of support to the rebellion can be re-expressed.

$$\begin{aligned} I(s) &= \left(1 - \frac{(\gamma+c)}{2\theta\phi}\right) \frac{1}{2\alpha\theta\phi} A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}(1+s)\right]^{\frac{\alpha}{1-\alpha}} \quad \text{when} \quad \theta > \frac{(\gamma+c)}{2\phi} \\ I(s) &= 0, \quad \text{otherwise} \end{aligned} \tag{8}$$

The minimum level of military forces necessary to deter rebellion presented in (5) is equivalent to the feasibility condition shown in (8). For the rebellion to be feasible, the rebel

[5] Unlike Grossman (1999), we do not introduce directly the army forces into the government's preferences. Despite its costs, having a large army should not necessarily affect negatively the government's preferences.

leader should be sufficiently efficient to overcome the coordination costs and the marginal benefit of informal activities. This condition is most easily met when the cost of deterrence is high. The government will also decide the equilibrium level of subsidy to be given to the worker. The government is constrained in his objective as the subsidy will need to be financed by a tax on production.

$$
\begin{aligned}
s = \frac{c+\gamma}{4\theta\phi} - (1-\alpha) \quad and \qquad & s \geq 0 \quad \text{if} \quad \theta \leq \frac{c+\gamma}{4\phi(1-\alpha)} \\
& s \geq 0 \quad \text{with} \quad I > 0 \quad \text{only if} \quad \alpha > \frac{1}{2} \\
\tau = \frac{\alpha+1}{2} - (1-\alpha)\frac{2\alpha\theta\phi}{c+\gamma} \quad and \quad & \tau \geq 0 \quad \text{if} \quad \theta \leq \frac{(\alpha+1)(c+\gamma)}{4\alpha(1-\alpha)\phi}
\end{aligned}
\tag{9}
$$

As shown in figure 2, the conditions under which the subsidy and the tax are positive, are met under the feasibility condition provided the marginal utility of labour is high enough ($\alpha > 1/2$). Otherwise, using a subsidy is a too costly strategy for the government. When θ belongs to the interval $[\frac{c+\gamma}{4\phi(1-\alpha)}, \frac{(1+\alpha)(c+\gamma)}{4\alpha\phi(1-\alpha)}[$, the government would only use the deterrence action. For θ above this interval, the government cannot resist anymore the uprising as it can use neither the soft nor the hard deterrence strategies and keep positive levels of production. This would be most likely to happen if the marginal utility of labour is very high. In the empirical work, we will consider the second interval where war is feasible ; no party is able to completely eradicate the other ; and the subsidy might be a valuable strategy to deter conflict.

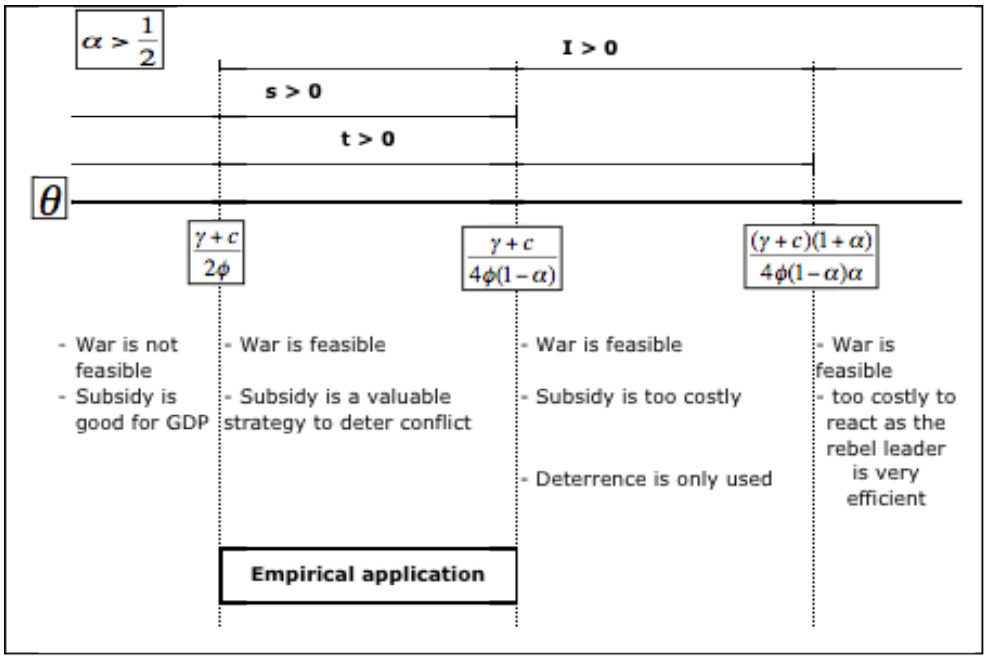
Given (9), condition (8) can be re-written.

$$
\begin{aligned}
I^* &= \left(1 - \frac{(\gamma+c)}{2\theta\phi}\right) \frac{A^{\frac{1}{1-\alpha}}}{2\alpha\theta\phi} \left[\frac{\alpha}{\gamma}\left(\frac{(c+\gamma)}{4\theta\phi} + \alpha\right)\right]^{\frac{\alpha}{1-\alpha}} \quad \text{when} \quad \theta > \frac{(\gamma+c)}{2\phi} \\
I^* &= 0, \quad \text{otherwise}
\end{aligned}
\tag{10}
$$

Computing w_l^*, L^*, Y^* and π^*, one can express the ratio of income between the local producers and the workers to capture a sense of inequality. A measure of absolute wealth per capita is obtained by dividing the equilibrium production by the numbers of workers.

$$
\begin{aligned}
\frac{\pi^*}{w_l^*} &= A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}\frac{(c+\gamma)}{4\theta\phi} + \alpha)\right]^{\frac{1}{1-\alpha}} \left[\frac{(1-\alpha)(\gamma+c)}{4\alpha\theta\phi} - \alpha\right] \\
\frac{Y^*}{L^*} &= \frac{1}{\alpha}\left[\frac{\alpha}{\gamma}\frac{(c+\gamma)}{4\theta\phi} + \alpha)\right]^{-1}
\end{aligned}
\tag{11}
$$

Figure 2: Range of parameters of interest



Therefore, condition (10) can be written in a more intuitive way.

$$I^* = \frac{1}{2\alpha\phi\theta}\left(\frac{\pi^*}{w_l^*}\right)\left(\frac{Y^*}{L^*}\right)\left(1 - \frac{(\gamma + c)}{2\theta\phi}\right)\left[\frac{(1-\alpha)(c+\gamma)}{4\alpha^2\theta\phi} - 1\right] \quad \text{when} \quad \theta > \frac{(\gamma + c)}{2\phi} \tag{12}$$

When the feasibility condition is met and if we restrict ourselves to the case where $s > 0$ (i.e. $\alpha > 1/2$), the equilibrium level of rebel forces is : [6]

- increasing in $\frac{\pi^*}{w_l^*}$. Increased inequality should be associated with more support to the rebellion. An increase in relative deprivation reduces the opportunity cost associated of joining a rebellion and hence raises the risk of local conflict.
- increasing in $\frac{Y^*}{L^*}$. Contrary to common wisdom, when controlling for the inequality level, an increase in income per capita could increase the support to the rebellion. This results from the fact that the opportunity effect being embodied by the inequality term, the income per capita only captures the looting effect. In other words, the larger is the production to loot, the keener the rebel leader is to hire fighters.

[6] The last term is interesting as it could be seen as a corrective term on GDP, reflecting whether the government is benevolent or not. A framework where the government is corrupted and only concerned about the net income of its clientele is given in appendix. In such a case, the subsidy is never a valuable strategy when the war is feasible.

- decreasing in c and γ. An increase in the coordination costs of organizing the uprising or of the marginal utility of informal activities would also decrease the equilibrium level of rebel forces.
- increasing in θ (when θ satisfied the feasibility condition). In other words, the relative efficiency increases the rebel forces. The same relationship holds for ϕ. However, when ϕ is too low, no rebellion can be organized as using deterrence is a very cheap strategy for the government. An increase in the cost of the deterrence action satisfying the feasibility condition, increases the incentives to expand the rebellion.

3 Empirical analysis

3.1 Conflicts in South Mexico

On January 1, 1994, *Chiapas* attracted international attention when the *Ejercicío Zapatista de Liberación Nacional* (*EZLN*) chose the date of the launch of the North American Free Trade Agreement (NAFTA) to occupy seven cities in this Mexican State. A ten-day confrontation occurred between the small group of rebels and the Mexican army, resulting in more than one hundred *Zapatistas* dead. However, this apparently short length of 10 days should not underestimate the conflict. On the one hand, the violence existed long before this event and went on even after the peak of the conflict in 1994. Particularly, the low-intensity but highly effective warfare conducted against the *Zapatistas* makes less clear the conventional distinction between peace and war times. Casanova (1996, 281) indicates that "from 1974 to 1987, 982 leaders were assassinated; 1,084 peasants arrested without legal cause; 379 seriously wounded; 505 kidnapped or tortured; 334 disappeared; 38 women raped, ...". After 1994, the violence continued. In 1995, President Zedillo ordered a new military offensive in order to arrest the alleged *EZLN* leaders. The offensive failed but was the source of a new wave of violence and forced displacement against indigenous people. In 1997, 45 *Zapatistas* were massacred in *Acteal* by paramilitaries. On the other hand, reducing the conflict to a period of ten days might give birth to misleading interpretations emphasizing the irrational nature of these violent events. [7] We will enlarge our analysis to two other states. *Guerrero* also hosted revolutionary movements through the emergence in the mid-nineties of the *Ejercicío Popular Revolucionarío* (*EPR*). No significant uprising is known in the third state, *Puebla*.

South Mexico and *Chiapas* are interesting cases for studying the role of inequality in conflict. Inequality appears to be the usual suspect but not necessarily, for objective reasons. Due to the symbolic synchronization with the launch of the NAFTA, many authors and leaders of the rebellion themselves have argued that the conflict in *Chiapas* was not just about this state but also against the whole neoliberal project, source

[7] Promoting an approach of continuity in the study of conflict, authors such as Galtung (1996) or Keen (2000) underline that the distinction between peace and war times is not so clear, as processes of exploitation and marginalization already existing in normal times are just reinforced in war times.

of inequalities. Since Benjamin (1989) and his book "A Rich Land, a Poor People", inequality between *Chiapas* and the rest of Mexico and a subsequent feeling of injustice have often been pointed out as a source of social discontent and rebellion (Duterme, 1998; Korsbaek, 1994; Morton, 2002). The same argument has been advanced as far as *Guerrero* is concerned. The contrast between high-value natural resources and poverty is unfortunately right. However, the short cut rich land-poor people as a source of conflict seems pretty weak. First, state inequalities in Mexico have historically been very high. GDP per capita in *Chiapas* and *Guerrero* were at least three times lower than the richest states, *Mexico City* and *Nuevo Leon* in 1970. After 30 years, this ratio has even doubled in 2000. The gap between some Southern states and the rest of the country has clearly widened since the mid-1980s, beginning of the entry into GATT, not only in terms of income per capita but also in terms of education achievement, life expectancy, etc. [8] However, this relative correlation is not a guarantee of causal inference. Put broadly, *Chiapas* and *Guerrero* were poorer than the rest of the country in 1970 and its inhabitants are still poorer now. The observed correlation does not explain why conflict occurs there and not in other poor states such as *Oaxaca*, *Tlaxcala*, *Veracruz-Llave*, *Tabasco*, *Michoacan* or *Puebla*. Second, it does not really explain why people decided to enter into conflict. As explained in section 3.4.2, *Chiapas* and *Guerrero* are far too heterogeneous to form one group in itself and overcome collective action problems. It is of utmost importance to understand individual motivations and explore the role of inequality at a more disaggregated level.

3.2 Methodology and data

A Probit model is constructed to estimate the effects of the explanatory variables on the probability of people to support or not the rebellion. The model to be estimated is based on the conditional probability of y given a vector of regressors x:

$$P(y = 1|x) = \Phi(\beta_0 + \beta_1 INEQUALITY + \beta_3 INCOME + \beta_3 ETHNICITY$$
$$+\beta_4 SIZE + \beta_5 DENSITY + \beta_6 CONTROL + \varepsilon)$$

where Φ is the cumulative density function of the standard normal distribution. Under the standard assumption of normality of the error terms (ε), coefficients are estimated by conditional maximum likelihood. Given the limits of cross-country analysis, a particular focus is given to local factors in understanding conflict in South Mexico. Data were collected at the municipal level on the basis of population census from the *Instituto Nacional de Estadística Geografía y Informatíca* (*INEGI*). The municipality is the smallest administrative unit in Mexico and hosts an average of 25,000 habitants in our sample. Data from 1970 are collected to introduce instrumental variables. To improve robustness in our analysis and as illustrated in figure 3, the sample is extended to two other states, *Guerrero* and *Puebla*. Rebel movements are recorded in

[8] This has been widely studied by notably Cikurel (2002), Rodríguez-Pose and Reaza (2002) and Tamayo-Flores (2001).

the former and not in the later. The sample is composed of 402 municipalities. As can be seen in table 1, these Southern states share common characteristics, being among the poorest states in Mexico and hosting a high percentage of people speaking an indigenous language.

Table 1: Three similar states

	Income per capita ($pesos)	People speaking an indigenous languages (%)
Mexico (country)	11338	7.49
Chiapas (110 muni)	6771	26.42
Guerrero (74 muni)	7572	13.40
Puebla (217 muni)	8661	14.11

Source: INEGI, Instituto Nacional de Estadística Geografía y Informática, data for 1990

Figure 3: *Chiapas*, *Guerrero* and *Puebla*

Note: The dark hatching indicates states out of the sample. Light hatching describes the states of which municipalities have been used for computing the index of inequality spillovers (see footnote 12)

Source: Geographic coordinates provided by CIMMYT (International Maize and Wheat Improvement Center)

3.2.1 Dependent variable : support to the rebellion

Working with so small geographical units, a difficult task is the determination of the dependent variable, i.e. whether or not a municipality is defined as supportive to the rebellion. For *Chiapas*, two indicators are used. On the one hand, Sonnleitner (2001) by studying in details the way the *EZLN* injunction to vote for the *Partido de la Revolucíon Democratíco* (*PRD*) candidate in 1994 and not to vote between 1995 and

1997 was followed, identifies 18 municipalities with a strong concentration of *Zapatista* supporters. On the other hand, Trejo (2002) investigates to which extend *Zapatista* Autonomous Municipalities (*Municipios Autónomos Zapatistas*, *MAZ*) have been created between 1994 and 1997. These *MAZ* are "de facto local governments that claim legitimate jurisdiction over newly defined territories within constitutionally elected mayors, governors and the President of Mexico and declare themselves only accountable to their *Zapatista* base communities and to the EZLN commanders in chief" (Trejo, 2002, 7). This author calculates a strong correlation between these rebel jurisdictions and the levels of violence that occur at the local level. This second indicator completes the sub-sample of 'supportive' municipalities with an additional unit, *Titalá*.

For *Guerrero* also hosting social unrest conducted by the *EPR*, the dependent variable was defined on the basis of the events of confrontations, murders and rapes exacted by the army soldiers against 'supposed' rebels and reported as such by the local press between 1995 and 1998, the peak of the conflict in that state (SIPAZ, 2005). The method is based on the assumptions that rebels are hidden where they find enough supporters and that exactions from soldiers make the population more likely to support the *EPR*. Eleven municipalities are determined as 'supportive' out of 74. No significant uprising is known in *Puebla* so that we end up with a sample of 402 municipalities, of which 30 are considered 'supportive' to local rebels. Those municipalities are represented in figure 4.

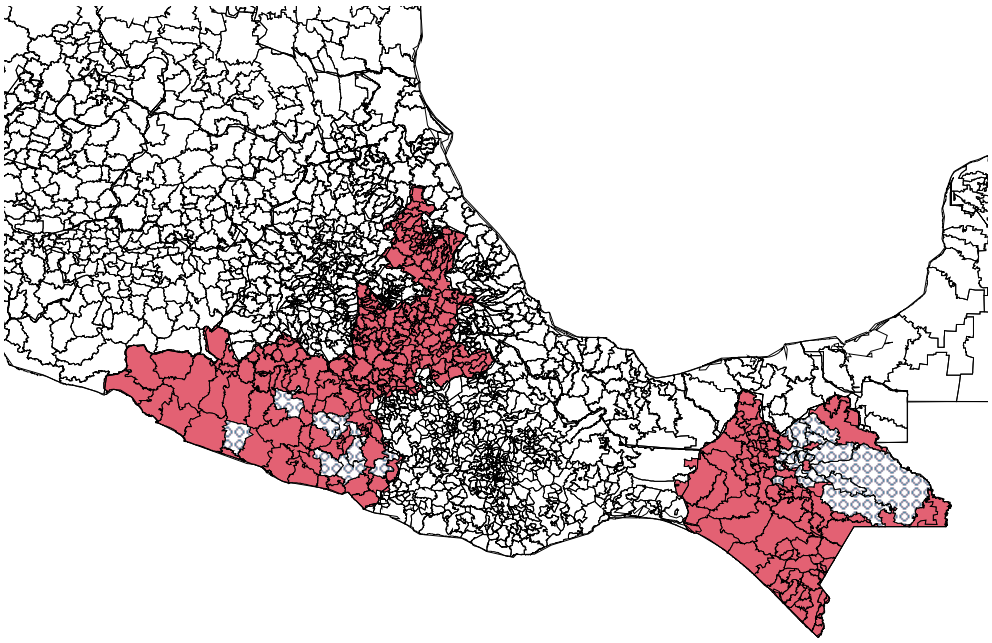
3.2.2 Explanatory variables

We follow the equilibrium condition found in (12) to identify and interpret our main explanatory variables. We assume the conditions are such that the war is feasible and the marginal utility of labour is sufficiently high to make labour subsidy a valuable strategy to deter conflict. In such a case, the following explanatory variables are defined :

Inequality : We should expect habitants of unequal municipalities to be more supportive to the rebellion. However, it is only true provided we control at the same time for the looting effect captured by the income per capita. The Gini coefficient and the polarization index of Esteban and Ray (1994) were estimated for each municipality. The later is argued to be much more significant in the study of conflicts. As argued by Esteban and Ray (1994, 820), "the phenomenon of polarization is closely linked to the generation of tension, to the possibilities of articulated rebellion and revolt, and to the existence of social unrest in general".

Income per capita : Condition (12) indicates that conditional on other variables, the level of production per capita should have a positive effect on the support to the rebellion. Contrary to Collier and Hoeffler (2004), an increase of the income per capita could therefore boost the potential loot the rebel leader could get from warfare (when controlling for the distributional effect of an increase in income per capita). The theoretical model underlines the necessity to study the effects of both income per

Figure 4: 402 municipalities, of which 30 are 'supportive'



Note: Municipalities in pink are those not assumed to be 'supportive' to the rebellion, while those with blue dots are assumed to be 'supportive'.

Source: Geographic coordinates provided by CIMMYT (International Maize and Wheat Improvement Center)

capita and the Gini coefficient simultaneously. A logarithm transformation has been applied to this variable.

Ethnicity as a coordination cost : Condition (12) indicates that the support to the rebellion is negatively correlated to the coordination costs. A standard approach of the coordination issue is measuring the ethnic dimension. According to our theoretical framework, ethnic diversity should make a rebellion more costly and hence less likely to occur. Collier (1998) also indicates that ethnic diversity should follow a non-monotonous relationship due to common problems of asymmetric information. Each individual has not the incentive to join the rebellion even if one shares one's ideas and is likely to free-ride on the others fighting. Furthermore, individuals have to fight before they can see the results of the struggle, which creates a time-consistency problem. Consequently, it is costly to coordinate the uprising. One way for leaders to overcome the problems of collective action is then to create a sufficiently high degree of trust by reinforcing the ethnic identity among potential fighters. In a very diverse society, identities are so numerous that it makes much harder to mobilize large numbers of people than in homogenous societies. Different measurements of diversity and polarization have been tested and are discussed in section 3.4.2. However, only the proportion of people of 5 years and more who speak an indigenous language is presented in the baseline regressions.

Density as a coordination cost : Coordination costs will also be approached with population density, transformed into logarithm. Contrary to Collier (1998) who uses population size as a proxy, we believe population density should better capture the difficulty to coordinate the rebellion. A denser population is assumed to ease the mobilisation of supporters by the rebel leader.

Size as a proxy for relative efficiency : Following our theoretical model, the rebel leader will have a strictly positive number of supporters provided $D < \hat{D}$, which corresponds to $\theta > \frac{(\gamma+c)}{2\phi}$ in (8) and (12). It is unfortunately impossible to have data on the relative efficiency of the rebels at the municipal level. However, we could first identify a general trend that could have made the uprising possible. Using data from Singer et al. (1972), we could observe a clear decreasing trend in the ratio of military expenditure over GDP between 1987 and 1991. Comparing to an average of 5% in the preceding 16 years, the ration decreases from 6.4% in 1987 to 2.14% in 1991. Second, the processes of trade openness and democratization since the beginning of the eighties made the deterrence action less coercive, suggesting that the sufficient condition to deter rebellion did not hold anymore after the end of the eighties. [9] Furthermore, the communication skills of the rebel leader *el subcommandante Marcos* has enabled the rebellion to gain some international support and increase the political cost for the Mexican government to 'nip the uprising in the bud'. Another dimension that could affect the relative efficiency of the rebels compared to the deterrence action of the central government is the presence of some geographic features, known in the literature as providing military advantages to rebels. Given the *Lacandone Selva* in *Chiapas*, such geographical dimension is likely to play an important role. Unfortunatly, no quantitative data could be found. However, comparing different maps, it appears that the presence of forests is associated with large municipalities. Therefore, we will control for the size of each municipality to capture the effect of geographical advantages for the rebellion. Such a control variable, transformed into logarithm, is therefore expected to have a positive effect on the probability to support the rebellion. However, one should note that this variable is likely to capture even more than this geographical advantage. [10]

3.2.3 Control variables

We also introduce control variables to test the robustness of our results in section 3.4.2. Such control variables and other qualitative information will allow us to test

[9] Empirically, democracy is found to have a non-monotonic effect on the risk of conflict. Evidence from Ellingsten (2000), Hegre et al. (2001), Reynal-Querol (2002) and Sambanis (2001) suggests partly democratic regimes are more prone to violent conflicts than either highly democratic or highly autocratic societies.

[10] As a negative proxy for the presence of forests, we test the proportion of arable land. The average altitude of each municipality computed on the basis of locality data has been introduced as a proxy for the strategic advantage that could offer the presence of mountains. Without changing the results presenting in section 3.4, both variables are far from significant in all our specifications.

our results against alternative hypothesis. Most variables are computed from *INEGI* data for the year 1990:
- The literacy rate (Education) is the proportion of people of 15 years and more who can write and read.
- Primary is the proportion of people who have completed primary education.
- Secondary is the proportion of people who have completed secondary education.
- An infrastructure index is computed by averaging three ratios i.e. the proportion of people with access to electricity, to running water and to drainage services.
- Migration will be measured by either the proportion of people born in another country or living five years before in another country.
- The importance of the agricultural dimension is measured by either the proportion of land or labour used in agricultural or forestry activities. The proportion of arable land is also tested.

3.3 Dealing with endogeneity

A key issue in our empirical work is the treatment of the potential endogeneity problem. This is particularly true with the *Inequality* variable. We suspect the level of inequality to be correlated with some unobserved heterogeneity across municipalities, such as the quality of the local political system and notably its openness to minority groups. On the one hand, Reynal-Querol (2005) suggests that the inclusiveness of the political systems has a significant impact on the risk of conflict. If such inclusiveness is correlated with our *Inequality* variables, the results will be biased. On the other hand, this might be highly relevant in our case. Local governance has been recognized as playing a very determinant role in some municipalities within the state of *Puebla* (Ward and Rodríguez, 1999; Vanderbush, 1999). The adopted strategy is to find valid instruments, test them with the two-step procedure of Rivers and Vuong (1988) and use conditional maximum likelihood method of estimation to cope with the potential endogeneity problem.

Finding a valid instrument is not an easy task. A natural candidate is the lagged value of the *Inequality* coefficient from 1970. Income inequality in 1970 is strongly correlated with the Gini coefficient of 1990 and uncorrelated with the dependent variable of 1990. [11] To be a valid instrument, we could assume that the Gini coefficient of 1970 is not correlated with the unobserved local governance of 1990, as enough political cycles and changes in political representativeness should have occurred in the 20-year period. However, we cannot be certain that such an assumption will necessarily hold. Despite the huge political changes known in Mexico over that period, it might well be the case that the process of democratization has not reached some parts of South Mexico. Twenty years might not be long enough to ensure sufficient changes in the political inclusiveness in some municipalities of our sample. Two other instrumental

[11] The risk of weak instrument should be low as the coefficient of the lagged Gini coefficient in the linear projection of the Gini coefficient on the exogenous variables of the baseline regression is highly significant (0.406***).

variables will be introduced. First, we construct a measure of inequality spillovers by multiplying the Gini coefficient of each municipality by a weighting matrix whose weight decreases with distance while is nul for the concerned municipality. [12] Such a computation should reflect the way inequality in surrounding municipalities will affect the level of inequality in the concerned municipality. Such an index is significatively correlated with the level of inequality of the concerned municipality but is unlikely to hold any relationship with the local governance, being completely exogeneously determined. Furthermore, it does not appear to affect the dependent variable by other channels than the Gini coefficient. The second alternative instrument is given by the proportion of people not wearing shoes in 1970. Such an indicator was collected as a critical measure of relative deprivation in 1970 by *INEGI* and is strongly correlated with the Gini coefficient of 1990. However, due to a general improvement between 1970 and 1990 with regards to the wearing of shoes, the relationship with the local governance of 1990 has fade away overtime.

Similar problem of endogeneity might be raised with the introduction of the income per capita. An equivalent strategy is adopted. We will introduce not only the lagged value of this indicator as an instrumental variable but also another alternative, the closest distance to the US. In the tradition of geographical economics, the distance to a larger market should have a negative impact on the income per capita. Not surprisingly, the distance to the most Southern location of Texas, ranging between 500 and 1300 kilometers by air (and up to 1600 kilometers by land) is highly significant in the linear projection of the income per capita on other variables. Exogenously determined, such a geographical variable is unlikely to be correlated with the local governance of the municipalities.

3.4 Results

3.4.1 Does inequality make us rebel?

Contrary to results obtained from cross-country analysis, income inequality measured by the Gini coefficient at municipal level is significant in motivating people to support the rebellion. The Gini coefficient in regression (1) of table 2 is positive and highly significant. The corresponding average partial effect accounts for 0.013. Regression (2) tests the income per capita apart from the Gini coefficient. Like Collier and Hoeffler (2004), we found a negative coefficient. [13] Being negatively correlated, both results are likely to be downward biased by the omision of one or the other variable. When

[12] To avoid the bias introduced by the discountinuity of space in our sample, we add 17 other states in South Mexico (the lower half of the country in light hatching in figure 3), i.e. covering about 1825 municipalities and 25 million people to compute the inequality spillovers and then, select the subsample corresponding to our 402 municipalities.

[13] Introducing the tax revenue (excluding federal transfers) by municipality, as an alternative proxy for the potential loot the rebel leader can get from warfare, is insignificant. Nevertheless, it does not alter significantly the results of other variables in the different specifications presented in this section.

introduced in the same regression (3), both coefficients are likely to better distinguish the opportunity cost effect from the looting one but become insignificant. However, we might suspect these variables to be correlated with the unobserved quality of local governance. [14] Using the lagged Gini coefficient, the inequality spillovers and the proportion of people not wearing shoes in 1970 as valid instruments, the Rivers and Vuong (1988) test confirms our intuition that the Gini coefficient is likely to be endogenous. Instrumentalizing the Gini coefficient increases the effect of inequality on the probability to support the rebellion. Regressions (4), (5) and (6) give a coefficient between 13.22 and 13.74 and an average partial effect between 1.5 and 4.5. Regression (7) confirms our results when using all instrumental variables. [15] As far as the income per capita is concerned, the Rivers and Vuong (1988) test gives mixed results depending on the instrument adopted. As a matter of comparison, we could note that instrumenting the income per capita in regression (2) would increase the magnitude of the effect (-1.614*** using the lagged income and -2.345*** using the shortest distance to the US). However, it might well be the case that the endogeneity bias is magnified due to the non-inclusion of the Gini coefficient as an explanatory variable. Regression (8) gives the results when both the Gini coefficient and the income per capita are instrumented. The coefficient of the income per capita is closed to the one obtained when it is introduced as an exogenous variable. One technical difference is that it makes the maximization of the conditional log-likelihood function much more complex. Including other explanatory variables breaks the convergence procedure down. We will adopt the specification of regression (7) in the rest of the paper. [16] Cautiousness in the interpretation of the coefficient of the income per capita is required. The positive sign obtained, when controlling for the Gini coefficient should not be seen as a counter-evidence to most cross-country findings that war is more likely to occur in poor countries. Nevertheless, it could suggest that within these poor countries, the way wealth is spatially distributed may matter as far as conflict prevention is concerned.

Results presented in this section has been obtained when ethnicity is simply proxied by the percentage of people speaking an indigenous language. In line with the theory

[14] The risk of approximate multicollinearity is pretty weak. The Gini coefficient is negatively correlated with the income per capita with a coefficient of -0.61 (-0.76 when the second variable is transformed into logarithm).

[15] The overidentifying restriction test sheds some doubts about our choice of instrumental variables. However, this test is an approximation in the case of non-linear estimations such as our Probit specifications.

[16] Similar results are obtained in regression (9) where no logarithm transformation for the income per capita, the size and the population density. As verified for all specifications, not using this transformation only decreases the magnitude of the de-logged variables and slightly improve the significance of the 'ethnicity' variables. All results are also consistent with the results obtained under the Linear Probability Model estimation. Although it is not well specified as about one third of probability responses stand outside the unit interval, it provides an interesting starting point as estimated coefficients should be consistent with the average partial effects obtained by Probit estimations.

Table 2: Baseline regressions

Variables	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit
	(1)	(2)	(3)	(4)(IV)	(5)(IV)	(6)(IV)	(7)(IV)	(8)(IV)	(9)(idem 7)
IV				Gini70	Spillovers	Noshoes	Multiple	Spill + DistUS	No Logs
Gini	4.319***		4.063[0.125]	13.573***	13.222***	13.738***	13.445***	8.386***	12.478***
	(1.482)		(2.649)	(1.722)	(0.73)	(0.669)	(0.924)	(.)	(0.732)
APE	0.013		0.012	1.479	3.941	4.455	2.929	2.947	3.080
Income per		-0.781**	-0.064	1.955***	2.167***	2.251***	2.143***	2.831***	0.001***
capita (log)		(0.307)	(0.552)	(0.486)	(0.191)	(0.209)	(0.231)	(0.009)	(0.000)
APE		-0.003	-0.000	0.213	0.646	0.73	0.467	0.995	0.000
Ethnicity	5.769**	5.336**	5.733**	3.376*	2.127**	1.655	2.809**	1.885***	3.531***
	(2.371)	(2.339)	(2.389)	(2.018)	(1.061)	(1.582)	(1.19)	(0.504)	(1.144)
APE	0.016	0.02	0.017	0.368	0.634	0.537	0.612	0.663	0.872
Ethnicity^2	-2.694	-2.161	-2.684	-2.215[0.193]	-1.653[0.143]	-1.39	-2.047[0.106]	0.287	-3.603***
	(2.625)	(2.554)	(2.625)	(1.701)	(1.128)	(1.344)	(1.265)	(0.608)	(1.233)
APE	-0.008	-0.008	-0.008	-0.241	-0.493	-0.451	-0.446	0.1	-0.889
Size(log)	6.655***	7.204***	6.69***	1.747	-0.169	-0.463	0.62		-0.000
	(1.422)	(1.425)	(1.455)	(2.192)	(0.808)	(1.569)	(0.935)		(0.000)
APE	0.019	0.027	0.2	0.19	-0.05	-0.15	0.135		-0.000
Density(log)	0.804***	0.872***	0.812***	0.234	-0.03	-0.055	0.057		-0.001***
	(0.209)	(0.219)	(0.221)	(0.278)	(0.117)	(0.208)	(0.137)		(0.000)
APE	0.002	0.003	0.002	0.025	-0.009	-0.018	0.012		-0.000
Obs.	402	402	402	402	402	402	402	402	402
Log lik.	-44.403	-45.622	-44.396	430.3	434.43	426.24	442.5	146.22	417.25
Pseudo R^2	0.58	0.57	0.58						

Source: Standard errors are displayed in parentheses under the coefficient estimates. *, **, *** denote a coefficient estimate significantly different from zero at the 10%, 5% or 1% level. P-value closed to 0.1 is given between brackets. GINI refers to the Gini coefficient . APE means average partial effect. Gini70, Spill, Noshoes and DistUS respectively refer to the Gini coefficient of 1970, the index of inequality spillovers, the proportion of people not wearing shoes in 1970 and the closest distance to the US.

considering this variable as a coordination cost, ethnic diversity increases the risk of conflict. Nevertheless, a too diverse society makes an uprising too costly to coordinate and explains the negative sign obtained for its square. Such a result is consistent with Collier (1998) who shows that highly diverse societies might be safer than less fractionalized societies, beyond a quite low level of diversity. We will discuss alternative measurements of the ethnic dimension. However, the significant effect found for the Gini coefficient is highly robust, as it will be unaltered in all alternative specifications. A denser population decreases the coordination cost and has a positive coefficient. In most cases, the size of the municipalities significantly increases the probability to support the rebellion. This variable could reflect the importance of physical geography such as the presence of forests, although it is also likely to capture other unobserved differences across municipalities.

Finally, although argued to be a much more relevant factor of conflict, the polarization index does not bear as much explanatory power as the Gini coefficient. [17] Assuming all variables exogeneous, the coefficient of the polarization index is insignificant. No significant coefficient is found when the same variables are estimated considering the polarization index as endogeneous. So poor results might be due to the unfavorable feature of the small size of municipalities. With small geographical units, polarization might matter less than the feeling of injustice engendered by overall inequality.

3.4.2 Robustness checks

Ethnicity is a difficult issue to handle in the literature on conflicts, as it appears to be malleable both over time and over space. Therefore, our measure of ethnicity as a proportion of people speaking an indigenous language may seem too simple. We identify in our sample about 37 linguistic groups such as the *Tzetzales*, the *Tzotziles*, the *Choles*, the *Tojolabales*, the *Zoques* and the *Mams*. [18] Such ethnic diversity would make rebellion a costly strategy to follow for the rebel leader. In our case, South Mexico appears to be ethnically very diverse. In addition to the 85% of the population studied who only speak Spanish, 15% of the population are composed of 36 groups whose members speak an indigenous language but which only represent between 0 and 6% of the sample, taken separately. However, the ethnic landscape is very much contrasted when looking at municipal data. For instance, people speaking *Tzetzal* represent 1.8% in our sample but 80.6% in *Chilón* (*Chiapas*); *Tzotzil*, 2.8% but 78.3% in *Chalchihuitán* (*Chiapas*); *Chol*, 1% but 61.48% in *Salto de Agua* (*Chiapas*), *Mexicano o nahuati*, 6.6% but 93% in *Zoquitlan* (*Puebla*) and *Mixteco*, 1% but

[17] This index was evaluated at the two particular values of α suggested by Esteban and Ray (1994) and reflecting different degrees of 'polarization sensitivity', i.e. $\alpha = 1$ and $\alpha = 1.6$. From Rivers and Vuong (1988) testing procedure and using the lagged variable from 1970, the polarization index does not seem to be affected by any endogenous bias.

[18] With the exception of the simple ratio of people of 5 years and more speaking an indigenous language, no detailed information has been found for the year 1990. Assuming that relative differences between municipalities in ethnic distinctions is stable over a 10-year period, we use data from 1980 to compute the various indexes of ethnic diversity and polarization.

83% in *Chigmecatitlán* (*Puebla*). Considering only the proportion of people speaking an indigenous language could also overlook a situation where two groups are polarized within a municipality. Therefore, we test the robustness of our results to several diversity and polarization indexes (details of computation are given in annex B). As illustrated in the first case in regression (1) of table 3, measurements of diversity such as the indexes of ethnic fragmentation (Taylor and Hudson, 1972), of peripheral diversity (Desmet et al., 2005), of maximum potential social fragmentation (Collier and Hoeffler, 2004) and a "Dominance" dummy indicating whether one linguistic group other than the only Spanish-speaking one is in majority do not significantly affect the probability to support the rebellion. As shown in regression (2) of table 3 for the most standard one, this is also the case for the polarization indexes of Esteban and Ray (1994) and Reynal-Querol (2001). With the sole exception of the index of peripheral polarization which gives the expected non-monotonic shape and similar results presented in regression (3) of table 3, the simple ratio of people speaking an indigenous language performs better than other more sophisticated indexes. This might seem puzzling but could have a straightforward explanation. [19] Many anthropologists have indeed witnessed the emergence of a shared identity as Indian, among ethnic groups such as the *Tzetzales*, the *Tzotziles* and the *Tojolabales* (Collier, 1995; Mattiace, 2001; Favre, 2002; Speed, 2002). The emergence of a common identity among Indians could have been a necessary condition for overcoming problems of asymmetric information in coordinating the uprising. [20] The anthropologist hypothesis reconciles the observed ethnic diversity with Collier's prediction. Indeed, high diversity within South Mexico can actually explain why social movements faced great difficulties to stabilize their membership. From the 1970s to the mid-1980s, many social organizations emerged, under the impulsion of the Bishop Samuel Ruiz García. These organizations had the common characteristics to be very locally based and to compete with each other (Favre, 2002). They were so numerous that Indians and peasants could join a formation with the expectation of quasi-immediate benefits and join another, otherwise. The only way to stabilize membership in such circumstances was to offer benefits from funds of public policies but with the obvious risk of becoming a political weapon of the dominant party, the *Partido Revolucionarío Institucional* (*PRI*). Consequently, no armed forces could have been organized due to collective action problems. In addition to the cost of deterrence action, what makes military mobilization possible could have been a re-definition of identity among Indians in South Mexico thanks to the important work of consciousness-raising, exercised locally by external actors such as Marxist activists, liberalization theologians and the pre-cited Bishop. Furthermore, all these alternative measurements of ethnicity strengthen the robustness of our results. Consistently with our theoretical framework, the Gini coefficient measured at

[19] In this case, the well-behaved form of the index of peripheral polarization could be due to the fact this index is actually closer to the simple ratio of people speaking an indigenous language.

[20] Although naturally bearing the function of anonymity, the wearing of mask can be seen as a clever way to represent an indigenous movement but sufficiently inclusive to avoid the coordination costs involved by a possible ethnic identification.

municipal level always significantly increases the probability to support the rebellion in South Mexico.

Alternative explanations have been found in the literature on conflicts in South Mexico. First, contrary to what has been claimed by Mexican officials about the instrumentalization of the conflict by migrants from other countries such as Guatemala, the proportion of people born in another country is far from significance in regression (4) of table 3. The same is true when the impact of migrants is estimated with the percentage of people that used to live in another country in 1985, i.e. five years before. Second, the unique measurement of income inequality might be seen as too restrictive to reflect the potential role of inequality in conflict. Income certainly matters as an indicator of relative deprivation, but might not be enough to capture the complexity of the decision that leads to support the rebellion. Unequal access to education might play an important role in motivating people to rebel and is proxied by the percentage of literate people in each municipality. Access to education, measured by the literacy rate, is not significant in regression (5). Similar outcome is obtained when access to education is proxied by the percentage of people with primary education. The non-significance of education might be explained by a general progress in literacy and primary schooling in previous decades. [21] On the contrary, the percentage of people with secondary education has a significant effect on the probability to support the rebellion. As indicated in regression (6), the sign of its coefficient is positive suggesting the more educated you are, the more prone to rebel you would be. A possible explanation could be the mismatch between improved skills and job opportunities. The lack of well-paid job opportunities could be a source of frustration for an increasingly skilled workforce. Such a factor of conflict is not captured by the theoretical model as unemployment and heterogeneity among workers are not considered. Although such a conjecture requires further theoretical and empirical investigations, a similar argument is theoretically supported by Azam (2005) in another context to explain motivates for suicide-bombings. An additional dimension for relative deprivation, i.e. access to infrastructure is proxied by averaging the proportions of households with access to electricity, running water and drainage services for each municipality. This constructed index significantly affects the probability to support the rebellion and has an expected negative sign. Some authors have also presented the *Chiapas* conflict as the results of a sectoral crisis. The agricultural and forestry sector would have been particularly harmed by trade reforms undertaken since the mid-eighties. However, regressions (8) and (9) indicate that the coefficients of agricultural and forestry land or the proportion of the population working in such activities are insignificant.

Finally, other factors could be missing in the analysis due to data unavailability. First, uneven land distribution has often been pointed out as responsible for protest in rural areas. South Mexico is no exception. Most of the literature presents the land issue as an opposition between a few big landowners who would own the majority of the land

[21] Bourguignon and Morrisson (1998) provide a similar explanation in their study of the differences of income distribution among 33 developing countries.

Table 3: Other control variables

Variables	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit	Probit
	(1)(IV)	(2)(IV)	(3)(IV)	(4)(IV)	(5)(IV)	(6)(IV)	(7)(IV)	(8)(IV)	(9)(IV)
Control	(a)	(b)	(c)	Migration	Literacy	Secondary	Infrastructure	Agri(ha)	Agri(Labor)
Gini	13.665***	13.57***	14.117***	13.531***	13.427***	13.663***	13.451***	13.441***	13.511***
	(0.851)	(0.958)	(0.865)	(0.929)	(0.927)	(0.923)	(1.003)	(0.98)	(0.864)
APE	3.459	3.211	3.451	2.96	2.889	3.128	2.583	2.965	3.212
Income per	2.109***	1.947***	2.177***	2.154***	2.159***	2.042***	2.492***	2.121***	2.188***
capita (log)	(0.220)	(1.947)	(0.213)	(0.228)	(0.234)	(0.133)	(0.265)	(0.247)	(0.219)
APE	0.534	0.461	0.532	0.471	0.464	0.467	0.478	0.468	0.52
Ethnicity	0.425	5.277	6.919***	2.785**	2.631**	3.133**	2.609**	2.677**	2.715**
	(1.332)	(6.699)	(2.069)	(1.179)	(1.204)	(1.226)	(1.216)	(1.21)	(0.219)
APE	0.108	1.249	1.692	0.609	0.566	0.717	0.501	0.591	0.645
$Ethnicity^2$	2.022	-3.331	-11.009***	-2.053[0.103]	-1.974[0.102]	-2.301*	-2.058[0.114]	-1.92[0.131]	-2.076*
	(2.003)	(38.78)	(3.197)	(1.257)	(1.268)	(1.272)	(1.302)	(1.273)	(1.233)
APE	0.512	-0.788	-2.692	-0.449	-0.425	-0.527	-0.395	-0.423	-0.493
Size(log)	0.173	0.438	0.178	0.647	0.632	0.512	0.925	0.614	0.469
	(0.734)	(0.814)	(0.753)	(0.933)	(0.954)	(0.929)	(0.991)	(1.005)	(0.891)
APE	0.044	0.104	0.043	0.141	0.136	0.117	0.178	0.135	0.111
Density(log)	0.039	0.052	0.007	0.055	0.065	0.023	0.129	0.08	0.041
	(0.757)	(0.133)	(0.119)	(0.135)	(0.142)	(0.137)	(0.146)	(0.139)	(0.131)
APE	0.01	0.012	0.002	0.012	0.014	0.005	0.025	0.018	0.01
Control				-2.071	0.363	8.562*	-2.329***	-1.062	0.277
				(6.844)	(0.84)	(4.766)	(0.854)	(1.567)	(0.416)
APE				-0.453	0.078	1.96	-0.447	-0.234	0.066
Obs.	402	402	402	402	402	402	402	402	402
Log lik.	446.29	435.11	442.254	442.793	442.762	445.065	449.403	443.799	442.823

Source: Standard errors are displayed in parentheses under the coefficient estimates. *, **, *** denote a coefficient estimate significantly different from zero at the 10%, 5% or 1% level. P-value closed to 0.1 is given between brackets. All estimations are given when three instrumental variables are used to correct for the potential endogeneity of the Gini coefficient. (a) Regression (1) includes the Ethnolinguistic Fractionization index (ELF) of Taylor and Hudson (1972) as an alternative for the *Ethnicity* variable. (b) Regression (2) incorporates the Polarization index of Esteban and Ray (1994) as an alternative for the *Ethnicity* variable.(c) Regression (3) introduces the index of peripheral polarization as an alternative for the *Ethnicity* variable.

at the expense of the whole Indian population. Furthermore, the 1992 amendment of article 27 of the Mexican constitution, aiming at attracting new private investment and withdrawing the peasant rights to claim land, is considered as a source of determinant recruitment for rebel movements. However, although the article 27 amendment has certainly eased the mobilization of rebels, this image of opposition between big landowners and the whole Indian population might have been relevant until the mid-eighties but not anymore in the nineties. It is true that private landowners received the support of the federal and military-lead state governments through forced migration, expropriation of land, division of populations and quasi immunity for the private guards employed by the cattle ranchers. However, after the eighties, there has been a multiplication of small private landowners but not a concentration of land in a few hands. Favre (2002) shows numerically that the average surface of biggest properties has diminished from 1970 to 1993 while the average surface of the smallest properties also diminished due to demographic pressure. Moreover, land owned collectively by the community seemed to have multiplied in sensitive areas of South Mexico. It does not mean that the land issue is not a crucial one but the problem is likely to be the non-respect of property rights, pointing to the local misgovernance rather than land distribution per se. Well, the way the potential endogeneity of the Gini coefficient is treated, should control for the unobserved effect of political exclusion. Naturally, the construction of an index incorporating, e.g. the ratios of minority groups in governmental bodies, civil services, the army and the police could better assess the partial effect of the political dimension. Second, trade of high-value commodities has often been central in financing war and economic agendas have often substituted political aims at the expense of the most vulnerable citizens. [22] However, the financing capacity of the rebellion is likely to be captured by the income per capita, reflecting the so-called looting effect of our theoretical model.

4 Conclusion

By extending the theoretical model of Grossman (1999), the present paper sheds light on the need to study both relative and absolute deprivation simultaneously in the analysis of conflict. On the one hand, income inequality measured by the Gini coefficient at municipal level has significatively affected the support to the rebellion in South Mexico. Therefore, cautiousness is required in the interpretation of the results obtained by (myopic) cross-country analysis that conclude that inequality does not play any role in motivating people to enter into armed conflicts. The literature has had huge implications on the way institutions such as the World Bank assess the risk of conflict and elaborate policies for conflict prevention and resolution. This paper

[22] No particular lootable resource has been identified in South Mexico. It might be the case that such a variable matters less than in other conflicts, such as the ones described in Africa. The uprising in South Mexico appears to be a 'cheap' fight to finance seeing the reported sources of financing the struggle. Harvey (1998, 167) indicates that weapons were mainly bought on the black market with "money that had previously been used for religious fiestas".

suggests, at least, that a one-fits-all approach to conflict could neglect economic inequality as an important factor. Contrary to Collier (2000)'s policy recommendations, reducing inequalities could well be effective in some cases to promote civil peace, basic condition for sustainable development. On the other hand, the theoretical framework suggests that a rise in income per capita could well exacerbate the risk of conflict, as it could increase the potential loot of the rebel leader. Such a result is supported by our empirical analysis of South Mexican municipalities. As a word of cautiousness, it does not contradict the general consensus obtained by cross-country analysis following which conflict is more likely in poor countries. However, it does suggest that the way growth is spatially distributed within countries under risk of conflict, does matter to prevent new deadly confrontations. This is certainly a promising path for further research. Finally, by verifying the robustness of the results to different specifications, the paper also emphasizes the complexity of ethnic identification. In the present case, the anthropologist hypothesis following which the emergence of a common identity across different ethno-linguistic groups was a necessary condition for group mobilization finds some empirical support. Nevertheless, the limited scope of this analysis constitutes an obvious constraint for being more affirmative on the issue.

References

Azam, J.-P. (1995). How to pay for the peace? a theoretical framework with references to african countries. *Public Choice*, 83:173–184.

Azam, J.-P. (2005). Suicide-bombings as inter-generational investment. *Public Choice*, 122:177–198.

Azam, J.-P. and Mesnard, A. (2003). Civil war and social contract. *Public Choice*, 115:455–475.

Benjamin, T. (1989). *A rich land, a poor people: Politics and Society in Modern Chiapas*. New York University Press, Alberqueque, N. M.

Billon, P. L. (2001). Angola's political economy of war: The role of oil and diamons 1975-2000. *African Affairs*, 100:55–80.

Bourguignon, F. and Morrisson, C. (1998). Inequality and development: The role of dualism. *Journal of Development Economics*, 57(2):233–257.

Casanova, G. P. (1996). Causes of the rebellion in chiapas. *Identities*, 3(1-2):269–290.

Champernowe, D. G. and Cowell, F. A. (1998). *Economic Inequality and income distribution*. Cambridge University Press.

Chen, S., Datt, G., and Ravallion, M. (1991). Povcal: A program for calculating poverty measures from grouped data. World Bank Policy Reserach Department.

Cikurel, C. D. (2002). Why mexico's regional income convergence broke down. In *Conference on Spatial Inequality in Latin America*. World Institute for Development Economic Research and Universidad de las Americas-Puebla, Mexico.

Collier, G. A. (1995). Restructuring ethnicity in chiapas and the world. *Indigenous Affairs*, 3:22–27.

Collier, P. (1998). The political economy of ethnicity. In Pleskovic, B. and Stiglitz, J. E., editors, *Annual World Bank Conference on Development Economics 1998*, pages 387–405. World Bank, Washington D.C.

Collier, P. (2000). Economic causes of civil conflict and their implications for policy. World Bank, Washington D.C.

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

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

Cramer, C. (2003). Does inequality cause conflict. *Journal of International Development*, 15(4):397–412.

Desmet, K., Ortuño-Ortín, I., and Weber, S. (2005). Peripheral diversity and redistribution. *CORE Discussion Paper*, 44:1–28.

Duterme, B. (1998). *Indiens et Zapatistes, Mythes et réalités d'une rébellion en sursis*. Luc Pire, Bruxelles.

Ellingsten, T. (2000). Colorful community and ethnic witches brew. *Journal of Conflict Resolution*, 44:228–249.

Esteban, J.-M. and Ray, D. (1994). On the measurement of polarization. *Econometrica*, 62(4):819–851.

Fairhead, J. (2000). The conflict over natural and environmental resources. In Nafziger, E. W., Stewart, F., and Väyrinen, R., editors, *War, hunger and displacement. The origins of humanitarian emergencies*, pages 147–178. Oxford University Press, New York.

Favre, H. (2002). Chiapas 1993: Intento de análisis de una situación de insurección. *Cuadernos Americanos*, 91:25–71.

Fearon, J. D. and Latin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90.

Galtung, J. (1996). *Peace by Peaceful Means: Peace and Conflict, Development and Civilization*. CA: Sage publications, London.

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

Grossman, H. I. (1999). Kleptocracy and revolutions. *Oxford Economic Papers*, 51:267–283.

Gurr, T. R. (1970). *Why men rebel?* Princeton University Press, Princeton.

Harvey, N. (1998). *The Chiapas Rebellion, the struggle for land and democracy*. Duke University Press, Durnham and London.

Hegre, H., Ellingsen, T., Gates, S., and Gledish, N. P. (2001). Towards a democratic civil peace? democracy, political change and civil war, 1816-1992. *American Political Science Review*, 95:33–48.

Hirshleifer, J. (1995). Theorizing about conflict. In Hartley, K. and Sandler, T., editors, *Handbook of Defense Economics*. Elsevier Science B.V., Amsterdam.

Hirshleifer, J. (2001). *The dark side of the force : Economic foundations of conflict theory*. Cambridge University Press, Cambridge.

Keen, D. (1993). The kurd in iraq : How safe is their haven now? *Save the Children Funds*. London.

Keen, D. (1994). *The benefits of famine. A political economy of famine and relief in Southwestern Sudan 1983-1989*. Princeton University Press, Princeton.

Keen, D. (2000). War and peace: What's the difference? *International Peacekeeping*, 7(4):1–22. Special issue: Managing armed conflicts in the 21st century.

Keen, D. (2001). Disqualifying grievance? a response to the collier/hoeffler model as an explanation for civil war. *COPED Conference*, June. School of Oriental and African Studies, London.

Korsbaek, L. (1994). The indigeneous rebellion in chiapas. *Indigenous Affairs*, 1:12–17.

Mattiace, S. L. (2001). Regional renegociations of space. tojolabal ethnic identity in las margaritas, chiapas. *Latin American Perspectives*, 28(2):73–97.

Morton, A. D. (2002). La resurrección del maíz : Globalization, resistance and the zapatistas. *Journal of International Studies*, 31(1):27–54.

Muller, E. N. (1985). Income inequality, regime represiveness, and political violence. *American Sociological Review*, 50(1):47–61.

Murshed, S. (2002). Conflict, civil war and underdevelopment: an introduction. *Journal of Peace Research*, 39(4):387–393.

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

Noh, S. J. (2002). Production, appropriation and income transfer. *Economic Inquiry*, 40(2):279–287.

Olsson, O. (2007). Conflict diamonds. *Journal of Development Economics*, 82:267–286.

Reynal-Querol, M. (2001). *Religious and ethnic conflict, political systems and growth*. PhD thesis, London School of Economics.

Reynal-Querol, M. (2002). Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution*, 46:29–54.

Reynal-Querol, M. (2005). Does democracy preempt civil wars? *European Journal of Political Economy*, 21:445–465.

Richards, P. (1996). *Fighting for the rain forest. War, youth and resources in Sierra Leone*. James Currey, Oxford.

Rivers, D. and Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics*, 39:347–366.

Rodríguez-Pose, A. and Reaza, J. S. (2002). Economic polarization through trade: The impact of trade liberalization on mexico's regional growth. In *Conference on Spatial Inequality and Development*. Cornell University-London School of Economics-World Institute for Development Economic Research, London.

Roemer, J. E. (1988). Rationalizing revolutionary ideology : A tale of lenin and the tsar. In Taylor, M., editor, *Rationality and Revolution*, pages 229–244. Cambridge University Press, New York.

Rubin, B. (2000). The political economy of war and peace in afghanistan. *World Development*, 28(10):1789–1803.

Russett, B. M. (1964). Inequality and instability: The relation of land tenure to politics. *World Politics*, 16(3):442–454.

Sambanis, N. (2001). Do ethnic and nonethnic civil wars have the same causes? a theoretical and empirical enquiry. *Journal of Conflict Resolution*, 45:259–282.

Sen, A. (1997). *On economic inequality*. Clarendon Press, Oxford.

Singer, J. D., Bremer, S., and Stuckey, J. (1972). Capabiliy distribution, uncertainty, and major power war, 1820-1965. In Russett, B., editor, *Peace, War and Numbers*, pages 19–48. Sage, Beverly Hills, basis for Correlates of War National Material Capacities Dataset 1986-2001.

SIPAZ (2005). Sipaz information. SIPAZ report X(2).

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

Sonnleitner, W. (2001). Chiapas: de la rebellion armée à l'alternance politique. *Problèmes d'Amérique Latine*, 41:21–43.

Speed, S. (2002). Global discourses on the local terrain. human rights and indigenous identity in chiapas. *Cultural Dynamics*, 14(2):205–228.

Stewart, F. (2000). Crisis prevention: Tackling horizontal inequalities. *Oxford Development Studies*, 28(3):245–262.

Tamayo-Flores, R. (2001). Mexico in the context of the north american integration: Major regional trends and performance of backward regions. *Journal of Latin American Studies*, 33:377–407.

Taylor, C. and Hudson, M. C. (1972). *The World Handbook of Political and Social Indicators*. CT: Yale University Press, New Haven.

Trejo, G. (2002). Redefining the territorial bases of power: Peasants, indians and guerilla warfare in chiapas, mexico. *International Journal of Multicultural Studies*, 4(1):1–29.

Turton, D. (1997). *War and Ethnicity: Global Connections and Local Violence. Studies on the Nature of War*, volume 2. University of Rochester Press, New York.

Vanderbush, W. (1999). Assessing democracy in puebla: The opposition takes charge of municipal government. *Journal of Interamerican Studies and World Affairs*, 41(2):1–27.

Ward, P. M. and Rodríguez, V. E. (1999). *New Federalism and State Government in Mexico*. US-Mexican, Report 9, Austin.

A Alternative specification of the theoretical model

An alternative specification would be a situation where the government is only concerned about the local producers, i.e. the local elite. The same sequential game is played and the government problem is written as follows:

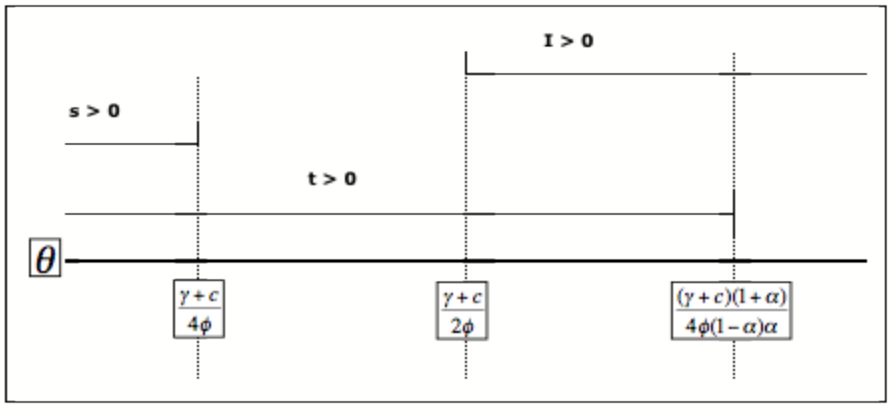
$$\begin{aligned} Q &= (1-\tau)\lambda Y - w_l L \\ \text{s.t.} \quad & \tau\lambda Y \geq s w_l L + \phi D \end{aligned}$$

As in the model of section 2, the government will decide on the levels of deterrence action, of taxes and subsidy.

$$\begin{aligned} D &= \frac{(\gamma+c)}{4\alpha\theta\phi^2} A^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{\gamma}(1+s)\right]^{\frac{\alpha}{1-\alpha}} \qquad (13) \\ s &= \frac{c+\gamma}{4\theta\phi} - 1 \quad \text{and} \quad s \geq 0 \quad \text{if} \quad \theta \leq \frac{c+\gamma}{4\phi} \\ \tau &= \frac{\alpha+1}{2} - (1-\alpha)\frac{2\alpha\theta\phi}{c+\gamma} \quad \text{and} \quad \tau \geq 0 \quad \text{if} \quad \theta \leq \frac{(\alpha+1)(c+\gamma)}{4\alpha(1-\alpha)\phi} \end{aligned}$$

The only change compared to the model presented in section 2 is the fact that the parameters for which the subsidy is positive, are more restrictive. Such a result is quite natural as the government is not benevolent anymore so that the social return on subsidy should be lower. In such a theoretical framework, the subsidy could never be a valuable strategy to deter war when the war is feasible. The hard stick will always be preferred to the soft manner when the local government is not benevolent.

Figure 5: Feasible intervals do not overlap



The same result is obtained when the government tax production rather than profits. Although this result might be insightful and raises the hypothesis that more corrupted government should never use a distributive policy as a strategy to deter rebellion, we do not consider this case as very insightful for our empirical work.

B Data sources and computation

Data come from different population census and the database provided by the the Instituto Nacional de Estadística Geografía y Informatíca (INEGI). To ensure a high degree of comparability between data from 1970 and 1990, a new municipality (San Juan Cancuc) in Chiapas has been incorporated into its former one (Tenejapa).

1. Gini coefficient: INEGI provides for each municipality the number of people by income groups, defined on the basis of the minimum salary (including no income). The Gini coefficient is calculated with grouped data so that no account of the inequality within the intervals of income is taken. As suggested by Chen et al. (1991), the midpoints were used for the closed intervals. Simulations indeed suggest that using midpoints to estimate Lorenz curves (from which the Gini coefficient is derived) does not create bias. For the open-ended interval at the top of the income distribution, the mean income for the interval was set at 30 percent above the lower bound. The lowest interval was set at 80 percent below the highest bound. Both values are recommended by Chen et al. (1991). Income intervals are defined in terms of minimum salary for 1990. With the sole exception of one municipality in the state of *Guerrero*, *Acapulco de Juarez* with a minimum salary of 10,309$ pesos on January 1, 1990, the minimum salary in the remaining 401 municipalities accounts for 8,405$ pesos. The formula provided by Champernowe and Cowell (1998) is used for computation:

$$Gini = \frac{1}{2n^2\bar{y}} \sum_{i=1}^{n} \sum_{j=1}^{n} |(y_i - y_j)| \tag{14}$$

   where y_i is the income attributed to individuals i ; n is the number of people living in the concerned municipality.

2. Income Polarization index: The same method has been used to deal with grouped data. Computation is based on Esteban and Ray (1994):

$$ER = K \sum_{i=1}^{n} \sum_{j=1}^{n} \Pi_i^{1+\alpha} \Pi_j |y_i - y_j| \tag{15}$$

   where Π_i is the proportion of the active population who declare to receive the income of the group i.

3. Income per capita is obtained by the ratio of the sum of the income of all the individuals (in the active population of the concerned municipality) divided by total population:

$$\text{Income per capita} = \frac{\sum_{i=1}^{n} Y_i \Pi_i}{n_k} \tag{16}$$

where Y_i is the income attributed to the group i

4. Baseline 'Ethnicity' variable: Proportion of people of 5 years old and more who speak an indigenous language.

5. The Ethnolinguistic Fractionization index (ELF) follows Taylor and Hudson (1972):

$$ELF = 1 - \sum_{i=1}^{n} \Pi_i^2 \tag{17}$$

where Π_i is the proportion of people that speak language i (religion i when applied to religious fractionization). This index reflects the probability that two randomly selected individuals within a municipality belong to the same linguistic group.

6. The Reynal-Querol index is naturally given by Reynal-Querol (2001):

$$RQ = 1 - \sum_{i=1}^{n} \left[\frac{0.5 - \Pi_i}{0.5} \right]^2 \Pi_i \tag{18}$$

7. The Peripheral Diversity/Polarization index follows Desmet et al. (2005):

$$PD(DOW) = \sum_{i=1}^{n} (\Pi_i^{1+\alpha} \tau_{oi} + \Pi_i) \Pi_o \tau_{oi} \tag{19}$$

where Π_o is the proportion of people only speaking the dominant language (Spanish in our case) ; Π_i is the proportion of people speaking the language i ; τ_{oi} is the distance between languages. When $\alpha < 0$, you obtain the index of peripheral diversity (PD), when $\alpha > 0$, you obtain the index of peripheral polarization (DOW).

These indexes should better reflect the dominant position of the only-Spanish-speaking people, not necessarily in terms of number (other linguistic groups are in a few cases in majority) but in terms of social, political and economic powers. This group is therefore assumed to be the dominant group in all municipalities. Ideally and as suggested by Desmet et al. (2005), the measure of distance would be based on the proportion of cognates between languages. To our knowledge, such information is unavailable for the different languages of our sample. Therefore, the 'distance' between languages was arbitrarily assumed to be equal to 0.5 between two indigenous languages and 1 between any of them and the Spanish language. The value 1 is in line with the distance suggested by Desmet et al. (2005) between Mayan and Spanish. The value 0.5 is chosen arbitrarily.

8. The Maximum Potential Social Fractionisation (MPSF) follows Collier (1998, footnote 8), i.e. the product of ethnic and religious diversity indexes, plus whichever index is the greater.

9. Control variables are given in section 3.2.3.

C Descriptive statistics

Table 4: Variables related to baseline regressions

Variables	Obs.	Mean	st.deviation	Minimum	Maximum
Supportive	402	0.0746	0.2631	0	1
Gini	402	0.6054	0.1261	0.3268	0.9533
Inc. per capita*	402	2035.164	1106.341	228.553	7115.858
Ethnicity	402	0.2158	0.2918	0.0012	0.8706
Size*	402	418.723	748.372	8.93	1061.6
Density*	402	115.633	204.386	2.603	2125.147

Note: * means that the variable has been transformed into logarithm before being introduced in regressions

Table 5: Control variables

Variables	Obs.	Mean	st.deviation	Minimum	Maximum
Ethnolinguistic Frac.	402	0.2054	0.1969	0.0018	0.6682
Polarization Index	402	0.0668	0.0558	0.0009	0.1648
Peripheral Polarization	402	0.1334	0.1724	0.0009	0.8698
Literacy rate	402	0.3206	0.1432	0.0619	0.8086
Primary	402	0.0899	0.0326	0.0181	0.1997
Secondary	402	0.0391	0.0234	0	0.2129
migration(born)	402	0.0023	0.0094	0	0.0969
migration(5 years)	402	0.0012	0.0034	0	0.0461
Infrastructure	402	0.3662	0.1564	0.007	0.8359