Greed and Grievance in Civil War

CSAE WPS/2002-01

Paul Collier
World Bank

and

Anke Hoeffler
Centre for the Study of African Economies
21 Winchester Rd
Oxford OX2 6NA
Tel: +44 (0)1865 274 554
anke.hoeffler@economics.ox.ac.uk
(corresponding author)

March 13, 2002

Abstract

We investigate the causes of civil war, using a new data set of wars during 1960-99. We test a `greed' theory focusing on the ability to finance rebellion, against a `grievance' theory focusing on ethnic and religious divisions, political repression and inequality. We find that greed considerably outperforms grievance. Consistent with the greed theory, both dependence upon primary commodity exports and a large diaspora substantially increase the risk of conflict. Inconsistent with the grievance theory, greater ethnic and religious diversity reduce the risk of conflict. The results are robust to correction for outliers, alternative variable definition, and variations in estimation method.

JEL Classification: C23, C25, D74, H56, O13
Keywords: Conflict, Development, Natural Resources, Panel Data

Non-Technical Summary

Civil wars are now far more common than interstate wars. During 1999 there were 27 major armed conflicts globally, all but two took place within national boundaries. In this paper we investigate the causes of civil war using a global data set including nations at war as well as peaceful countries.

According to popular perceptions grievances are often seen as the main causes of rebellion. However, we suggest that those factors which determine the financial and military viability of a rebellion are more important than objective grounds for grievance. In order to create and maintain a rebel organization the rebels have to be paid and military equipment has to be purchased. To test these alternative hypotheses we construct two competing models: a 'grievance' and a 'greed' model. The 'grievance' model examines inequality, political oppression, and ethnic and religious divisions as causes of conflict, while the 'greed' model focuses on the sources of finance of civil war.

We find little evidence for grievances as a determinant of conflict. Neither inequality nor political oppression increase the risk of conflict. However, we find some evidence that societies characterized by 'ethnic dominance,' i.e., where one ethnic group makes up 45-90 percent of the population, have a systematically higher risk of civil war. Our 'greed' model provides much better explanatory power. We are able to identify several sources of finance for rebellions. Income from natural resource predation such as diamonds in Angola, drugs in Columbia and timber in Cambodia are often quoted as important sources of finance for the rebel movements. Our empirical analysis confirms that countries with abundant natural resources have a higher risk of conflict. However, the relationship between natural resources and conflict risk is non-linear. Countries with a very high dependence on natural resource incomes have a relatively lower risk of conflict. We hypothesize that natural resources not only represent a source of rebel finance but also of government revenue. In poor countries governments often tax primary commodities at high rates and can use this income to strengthen the state. Furthermore, we find some evidence that countries with a large diaspora abroad experience higher conflict risks. Diasporas, such as for example the Kurdish community in Europe, the Lebanese in West-Africa and the Tamils in North America, often have the ability to use large financial resources and publicity to keep combatants active in their native countries.

In spite of the relatively low explanatory power of the 'grievance' model we cannot reject it in favor of the 'greed' model and thus combine the two models. In addition to the factors mentioned above we find that poorer countries are more likely to experience civil wars. The opportunity cost for potential rebels are low and thus make recruitment easier. A high enrolment rate of males in secondary school also reduces the conflict risk. We also find that history matters. If a country experienced a conflict recently the risk of recurrent conflict is high, however, this risk falls proportionately to the length of the peace period. Sustained peace makes renewed rebellion less likely. Contrary to the 'grievance' theory, social fractionalization, measured as religious and ethnic diversity, *lowers* the risk of conflict. Typically rebel organizations recruit their members from similar backgrounds and diversity may make it more difficult to generate a large rebel force and to maintain cohesion during the war. Diversity would

thus reduce the risk of conflict. We also control for geographical characteristics since they are likely to influence combat strategy and thus the relative military capability of the regular and rebel forces. We find the risk of civil war to be higher in more mountainous nations and countries in which the population is unequally distributed.

Overall, these results are consistent with economic models of conflict risk in which the critical parameters are the financial opportunities for rebels, the social and geographic constraints which they face, and the financial capability of the government to provide defense and other public services. They are harder to reconcile with accounts of conflict which stress ethnic, religious, political or economic grievances.

1. Introduction[1]

Civil war effects many of the world's poorest countries. It is now far more common than international conflict: of the 27 major armed conflicts listed by the Stockholm International Peace Research Institute for 1999, all but two were internal.

In this paper we develop an econometric model which predicts the outbreak of civil conflict. In the process, we compare two contrasting accounts of the causes of conflict: preferences and constraints. The political science literature has stressed differences between societies in the preference for conflict: societies differ in the severity of objective grievances such as political repression or inequality. The economics literature has taken rebellion to be an economic activity with pay-offs akin to crime, and has focused on differences in constraints. In Section 2 we set out these contrasting accounts and discuss measurable variables which might enable us to test them. We present descriptive data on the 78 large civil conflicts which occurred between 1960 and 1999. In Section 3 we use non-nested tests to discriminate between the two accounts of conflict and develop an integrated model which provides a synthesis. Section 4 presents a range of robustness checks and Section 5 offers a concluding discussion of the results.

We conclude that differences in constraints are far more important in explaining the incidence of conflict than are differences in objective grounds for grievance. A particularly powerful risk factor is dependence upon primary commodity exports. A likely explanation is the scope these activities provide for extortion by rebel organizations. Whether such extortion directly motivates rebellion, or simply provides critical finance which facilitates the violent pursuit of other objectives, is beyond the scope of our paper.

2. Rebellion: Theories and Descriptive Evidence

The definition of a civil war which we use in this paper is of an internal conflict with at least 1,000 combat-related deaths, with both an identifiable rebel organization and government forces suffering at least five percent of these casualties. This definition has become standard following the seminal data collection of Singer and Small (1982, 1994). We use an expanded and updated version of their data set which covers 161 countries over the period 1960-99 and identifies 78 civil wars. Table 1 lists the civil wars.

[1] Previous versions of this paper have benefited from presentations at CERDI, NYU, LSE, Oxford, Princeton, Lisbon and the World Bank. We would like to thank participants for comments, especially Todd Sandler for helpful written suggestions.

Table 1: Outbreaks Of War

Country	Start of the War	End of the War	Previous War	GDP sample	Secondary Schooling Sample
Afghanistan	04/78	02/92			
Afghanistan	05/92	Ongoing	*		
Algeria	07/62	12/62	*		
Algeria	05/91	Ongoing	*	*	*
Angola	02/61	11/75			
Angola	11/75	05/91	*	*	*
Angola	09/92	Ongoing	*	*	*
Azerbaijan	04/91	10/94			
Bosnia	03/92	11/95			
Burma/Myanmar	68	10/80	*	*	*
Burma/Myanmar	02/83	07/95	*	*	*
Burundi	04/72	12/73		*	*
Burundi	08/88	08/88	*	*	*
Burundi	11/91	ongoing	*	*	*
Cambodia	03/70	10/91	*		
Chad	03/80	08/88		*	
China	01/67	09/68	*	*	
Columbia	04/84	ongoing	*	*	*
Congo	97	10/97		*	*
Cyprus	07/74	08/74		*	
Dominican Rep.	04/65	09/65		*	*
El Salvador	10/79	01/92		*	*
Ethiopia	07/74	05/91		*	*
Georgia	06/91	12/93			
Guatemala	07/66	07/72	*	*	*
Guatemala	03/78	03/84	*	*	*
Guinea-Bissau	12/62	12/74			
India	08/65	08/65	*	*	*
India	84	94	*	*	*
Indonesia	06/75	09/82	*	*	*
Iran	03/74	03/75		*	*
Iran	09/78	12/79	*	*	*
Iran	06/81	05/82	*	*	*
Iraq	09/61	11/63	*		
Iraq	07/74	03/75	*	*	*
Iraq	01/85	12/92	*	*	*
Jordan	09/71	09/71		*	
Laos	07/60	02/73	*		
Lebanon	05/75	09/92	*		
Liberia	12/89	11/91		*	
Liberia	10/92	11/96	*		
Morocco	10/75	11/89	*	*	*
Mozambique	10/64	11/75			
Mozambique	07/76	10/92	*	*	*
Nicaragua	10/78	07/79		*	*
Nicaragua	03/82	04/90			
Nigeria	01/66	01/70		*	*
Nigeria	12/80	08/84	*	*	*

Table 1 continued

Country	Start of the War	End of the War	Previous War	GDP sample	Secondary Schooling Sample
Pakistan	03/71	12/71		*	*
Pakistan	01/73	07/77			
Peru	03/82	12/96		*	*
Philippines	09/72	12/96	*	*	*
Romania	12/89	12/89		*	*
Russia	12/94	08/96			
Russia	09/99	Ongoing	*		
Rwanda	11/63	02/64			
Rwanda	10/90	07/94	*	*	*
Sierra Leone	03/91	11/96		*	*
Sierra Leone	05/97	07/99	*	*	
Somalia	04/82	05/88		*	*
Somalia	05/88	12/92	*	*	*
Sri Lanka	04/71	05/71		*	*
Sri Lanka	07/83	ongoing	*	*	*
Sudan	10/63	02/72			
Sudan	07/83	ongoing	*	*	*
Tajikistan	04/92	12/94			
Turkey	07/91	ongoing		*	*
Uganda	05/66	06/66		*	*
Uganda	10/80	04/88	*	*	*
Vietnam	01/60	04/75	*		
Yemen	05/90	10/94			
Yemen, Arab Rep.	11/62	09/69	*		
Yemen, People's Rep.	01/86	01/86	*		
Yugoslavia	04/90	01/92			
Yugoslavia	10/98	04/99	*		
Zaire/Dem. Rep. of Congo	07/60	09/65			
Zaire/Dem. Rep. of Congo	09/91	12/96	*	*	*
Zaire/Dem. Rep. of Congo	09/97	09/99	*	*	*
Zimbabwe	12/72	12/79		*	*

Note: Previous Wars include war starts 1945-94.

To explain the outbreak of civil war we need to discover the circumstances which favor rebellion. A helpful distinction is between preferences and constraints: societies can be prone to conflict either because preferences for rebellion are atypically strong, or, because constraints upon rebellion are atypically weak. This distinction has fairly precisely divided the large political science literature on the causes of conflict from the much smaller economics literature. Political scientists have focused upon grounds for grievance, such as ethnic hatred, political repression or inequality. Economists have usually abstracted from differences in preferences, treating the objective of rebellion as financial gain (Grossman, 1999; Collier, 2000). For example, Grossman states `the insurgents are indistinguishable from bandits or pirates' (p.269).

The assumption that rebellions are motivated by greed is merely a special case of the focus upon constraints. An alternative constraints-based theory is that of universal grievance: all countries might have groups with a sufficiently strong sense of grievance to wish to launch a rebellion, so that rebellions will occur where they are viable. Thus, it might be argued that both the Michigan Militia in the USA and the FARC in Colombia were established to address grievances through rebellion, but only in Colombia were the constraints upon rebellion weak enough to enable violence to escalate to the level of civil war.

To be viable, a rebel organization must survive militarily against the government army, and for this it needs manpower and equipment. In turn, these create the need for finance. Hence, both a greed theory and a universal grievance theory predict that the risk of rebellion is increasing in the opportunities for rebel finance. That the FARC is able to generate around $500m per year from cocaine may have become its rationale, or may simply facilitate its original objective. We will refer to the constraint-based theories of rebellion by the shorthand of `greed', to contrast them with preference-based theories, while recognizing that they do not necessarily literally imply that the motivation for rebellion is exclusively, or even primarily, financial.

Greed-Rebellion

Only large rebel organizations generate casualties on the scale which defines civil war. Typically, in civil wars rebel organizations have between 500 and 5,000 employees but can be much larger: UNITA, the rebel organization in Angola, at its peak had around 60,000. Rebel organizations combine this pool of largely unskilled labor with guns, and with this endowment must raise considerable finance to meet their expenses. We consider three sources of finance which are widely used by rebel organizations: extortion, donations from diasporas, and subventions from hostile governments.

We propose that the endowment of unskilled labor and guns which characterizes rebel organizations is particularly suited to raise finance through the extortion of primary commodity exports. Several high-profile examples, such as diamonds in Sierra Leone and Angola, timber in Cambodia, and cocaine in Colombia, give the proposition at least superficial plausibility. Our proposition can equivalently be interpreted in two ways. On the universal grievance interpretation rebellions need to finance themselves, and the extortion of primary commodity exports offers the best opportunity for financial viability. In the limit, only where there are such opportunities can rebel organizations escalate to the scale needed for civil war. On the literal greed interpretation the extortion of primary commodity exports will occur where it is profitable, and the organizations which perpetrate this extortion will need to take the form of a rebellion.

Even if rebel organizations are quasi-criminal (by choice or necessity), their niche in the crime market is specialized. Most crime is subject to scale diseconomies and so a large organization would be out-competed by smaller enterprises. Extortion normally offers some scale economies in violence and so is organized in larger enterprises than other crime (Konrad and Skaperdas, 1998). The extortion of primary commodities resembles other extortion, but has one distinctive feature. Whereas most criminal

extortion targets urban-based commerce, rebellions target the production or transportation of primary commodities, and these locations are usually rural. This affects the technology of defense. As with other crime, the returns on extortion can be decreased by defensive measures. However, whereas the most efficient defenders of commerce are probably detectives, primary commodities need defense of a large physical space best suited to an army. The high location-specific rents which characterize primary commodity exports not only attract rebel extortion but also induce governments to impose heavy taxation. Taxation gives governments both an incentive and a means to provide military defense. Faced with a military defense, a viable extortion racket itself needs considerable military power. Whereas a normal extortion racket merely needs sufficient force to menace its victims (and perhaps to defend itself from rivals), primary commodity extortion needs enough force to be able to survive in confrontation with the army. These greater scale economies of violence produce an industrial structure of fewer, larger organizations than in commercial extortion. The largest of these violent, quasi-criminal organizations will sometimes meet the criteria for civil war.

Suppose that the rebel organization depends for its finance on the extortion of primary commodity exports, while the government raises finance both from heavy taxation of primary exports and from lighter taxation of other income. Differences in the both the structure of income and its level now give rise to differential proneness to conflict. Consider, first, the structure of income. Evidently, countries without primary commodity exports do not offer opportunities for their extortion and so should have a low risk of civil war. However, the USA has abundant primary commodity exports and yet these did not provide extortion opportunities to the Michigan Militia. Clearly, one factor in the USA was that the government was sufficiently rich to provide an effective defense. The scale of rebel force necessary for survival against the military forces of the US government, while unavoidably exposing itself through operating a primary commodity extortion racket, is evidently so large that the required financing is prohibitive. More generally, the threshold of rebel force required for survival is increasing in the government's military expenditure. As a result, the viability of rebellion need not be continuously increasing in the endowment of primary commodity exports. Beyond some point, the increment in potential rebel revenue may be more than offset by the increased rebel expenditure needed to survive against augmented government forces.

Before investigating the effect of primary commodities formally, it is useful to review some descriptive statistics.[2] In Table 2, we measure the ratio of primary commodity exports to GDP for each of the 161 countries. As with our other variables, we measure at five year intervals starting in 1960 and ending in 1995. We then take the history of conflict in the subsequent five years, and compare those in which a conflict broke out with those which were conflict free. Those in which conflict broke out were on average slightly *less* dependent upon primary commodity exports than those which sustained peace: 15 percent versus 17 percent. However, the standard deviation for the conflict countries is only about half that of the countries which sustained peace. The countries in which conflict broke out tended to be grouped around average dependence, whereas the countries which sustained peace tended to have either markedly below-average or markedly above-average dependence.

[2] For details on the data used in the paper see the Appendix.

Controlling for the structure of income, an increase in the level of income is likely to favor the government since, as income rises, the share of income taken in taxation also tends to rise. As the government is strengthened financially relative to the rebels, the risk of conflict is likely to be reduced. In Table 2 we repeat the comparison, this time in terms of the level of income rather than its structure. There is a very marked difference: the countries in which conflict subsequently broke out had less than half the mean income of those which sustained peace ($1645 per capita versus $4219 per capita).

A second potential source of rebel finance is from diasporas living in developed countries: a well-documented example is the assistance provided to the Tamil Tiger organization of Sri Lanka by Tamils in north America but the phenomenon is more general (Angoustures and Pascal, 1996). There are several reasons to expect that diasporas would increase the provision of finance for rebellion. Diasporas are usually much richer than the population in their country of origin. They are better-placed for collective action: emigrants have a cultural incentive to create diaspora organizations which can then discipline free-riding. They do not suffer the consequences of the conflict. Diasporas may harbor historical grievances which the rebel organization can exploit. We measured the size of the diaspora from US Census data. Although this neglects diasporas living in other countries, it has the major advantage of greater uniformity in the aggregate: an immigrant to the USA is evidently not operating in the same legal, organizational or economic environment as an immigrant to France. We measure the importance of the diaspora by comparing the number of immigrants from a country currently resident in the USA, with the domestic population of the country. In the formal analysis we decompose the diaspora into that part induced by conflict and that which is exogenous to conflict, but here we simply consider the crude numbers. Superficially, the data does not appear to support the hypothesis that diasporas increase the risk of conflict: diasporas are substantially smaller in the countries in which conflict subsequently broke out.

A third potential source of rebel finance is from hostile governments. An example would be the role of the government of Southern Rhodesia in financing the Renamo rebellion in Mozambique. To investigate the proposition empirically, we need some proxy for the willingness of foreign governments to finance military opposition to the incumbent government. We use the contrast between the Cold War and post-Cold War periods as such a proxy. During the Cold War each great power tried to destabilize those countries allied to the opposing power. The incentive to support rebel movements was reduced once the Cold War ended and so, to the extent that this source of finance is important as a cause of conflict, we would expect a reduction in its incidence post-1989. We should note that this is a controversial prediction. For example, Kaplan (2000) argues that the end of the Cold War `lifted the lid' off previously suppressed conflict.

Table 2: Descriptive Statistics

	sample (n=1167)					*no civil war (n=1089)*					*civil war (n=78)*				
	mean	st.dev.	min	max	N	mean	st.dev.	min	max	N	mean	st.dev.	min	max	N
War starts	0.067				1167	0				1089	1				78
Primary commodity exports/GDP	0.168	0.189	0.002	2.139	1083	0.169	0.193	0.002	2.139	1011	0.149	0.114	0.005	0.505	72
GDP per capita (const. US$)	4061	4323	222	33946	1010	4219	4403	257	33946	948	1645	1353	222	5832	62
Diaspora	0.017	0.029	0	0.195	649	0.018	0.030	0	0.195	610	0.004	0.005	0	0.016	39
Male secondary schooling	43.42	30.86	0.3	147	959	44.39	30.98	0.3	147	893	30.3	26.1	1	102	66
GDP per capita growth t-1	1.617	3.827	-22.08	14.409	861	1.74	3.76	-22.08	14.41	807	-0.226	4.312	-10.664	7.093	54
Forest cover	31.11	23.63	0.02	96.76	1100	31.33	23.78	0.02	96.76	1029	27.81	21.18	0.44	81.01	71
Mountainous terrain	15.82	20.75	0	94.3	1167	15.17	20.35	0	94.3	1089	24.93	24.03	0	81	78
Geographic dispersion	0.571	0.229	0	0.971	1021	0.569	0.234	0	0.971	949	0.603	0.147	0.238	0.916	72
Population density	150	522	0.63	6218	1103	156	539	0.63	6218	1029	62	69	3.56	288	74
Population in urban areas	45.11	25.43	2.08	100	1136	46.00	25.56	2.08	100	1060	32.7	19.97	2.56	76.26	76
Ethnic fractionalization	39.57	28.48	0	93	1110	38.64	28.22	0	93	1027	52.63	29.06	4	90	73
Religious fractionalization	36.09	24.00	0	79	1140	35.98	23.95	0	79	1066	37.70	24.90	0	69	74
Ethnic polarization α−1.6	0.078	0.039	0.002	0.156	1110	0.077	0.040	0.002	0.156	1027	0.076	0.038	0.017	0.139	73
Ethnic dominance (45-90%)	0.465	0.499	0	1	1110	0.465	0.499	0	1	1027	0.452	0.501	0	1	73
Income inequality	40.61	10.27	17.83	63.18	718	40.58	10.27	17.83	63.18	672	41.00	10.45	23.38	62.3	46
Land inequality	0.641	0.168	0.227	0.980	748	0.641	0.170	0.227	0.980	701	0.631	0.131	0.366	0.911	78
Democracy	3.91	4.22	0	10	980	4.07	4.27	0	10	908	1.821	2.740	0	9	72
Previous war	0.208	0.406	0	1	1167	0.185	0.388	0	1	1089	0.538	0.501	0	1	78
Peace duration	326.6	163.4	1	592	1167	334.2	159.8	1	529	1089	221.0	177.4	1	592	78

Our focus has so far been on the differential scope for raising rebel revenue. We now turn to rebel expenditure, and consider differential costs. Superficially, it might appear that the rebels and the government face the same recruitment problem. However, this ignores an important difference. To initiate rebellion the rebel force must grow rapidly, whereas the government army can be presumed to be in steady state with a relatively low rate of turnover. This may make the rebel organization more sensitive to the current state of the labor market. Controlling for the level of per capita income, the tighter is the market for young male labor, the more difficult is the rebel recruitment problem. This is consistent with evidence that the faster the rate of growth the lower is the probability of unconstitutional political change (Alesina *et al*. 1996).[3] The Russian civil war of 1919-21 illustrates the effect of short term opportunity cost on the rebel recruitment problem. The Reds and the Whites, both rebel organizations, suffered massive desertion (the obverse of the recruitment problem) with some four million men quitting. The desertion rate was ten times higher in summer than in winter: the recruits being peasants, the opportunity cost of their labor was much higher at harvest time (Figgis, 1997). We measure the state of the labor market through two proxies: the rate of economic growth per capita and the proportion of young males enrolled in secondary education. To reduce problems of endogeneity, we measure these over the five year period prior to the one in which we assess the conflict history. From the descriptive statistics both variables appear to matter: the countries with subsequent conflicts had slower economic growth and lower school enrollment.

A further potential influence on rebel costs is the presence of accumulated physical, human and organizational capital. If a country has previously had a rebellion there will be a stock of guns, former rebels who know how to use them, and probably a persistent, if quiescent, rebel organization. Hence, the costs of renewed rebellion are likely to be lower than those of initial rebellion. In our sample, the conflict episodes were twice as likely in countries which had had a previous conflict since 1945 as in those which had been peaceful. However, this cannot necessarily be interpreted as indicating that conflict increases the risk of further conflict. The pattern may be due to fixed characteristics which make conflict more likely. A dummy variable for a previous history of conflict is interpretable either as a proxy for accumulated rebellion-specific capital or as a proxy for omitted fixed effects. However, since capital decays this provides a way of distinguishing between it and fixed effects. We introduce a variable measuring the time since the previous conflict (with a maximum back to 1945). On average, conflict episodes are preceded by a much shorter period of peace.

In addition to revenues and costs, a further observable factor which can be expected systematically to affect the military feasibility of rebellion is geography. Forest cover may provide rebels with a safe haven. American efforts at defoliation in Vietnam suggest that this is at least perceived as militarily important. We investigate this

[3] The economic growth literature concentrates on the analysis of political instability as a determinant of economic growth (see for example Barro 1991, 1997). Alesina et al (1996) estimate a simultaneous equation system of economic growth and political instability. They present support for the hypothesis that political instability reduces growth. Lower growth does not seem to cause political instability, defined as the number of government changes. However, when they define political instability more narrowly as unconstitutional government changes they find that lower growth rates are a causal factor of political instability.

through the FAO measure of the proportion of a country's terrain which is covered in forest. The descriptive statistics presented in Table 2 show that war countries actually had a slightly *lower* forest coverage (29 percent) than peaceful countries (31 percent). Similarly, mountainous terrain may provide a safe haven. We could find no equivalent data set on the proportion of a country's terrain which is mountainous. Crude measures such as altitude tend to misclassify both plateaus and rugged uplands. We therefore commissioned a new index from John Gerrard, a physical geographer specialized in mountainous terrain. The descriptive statistics suggest that mountainous terrain is important: in civil war countries 25 percent of their terrain is mountainous, while only 15 percent of the terrain of the average peaceful country consists of mountainous terrain. Geographic dispersion of the population may make military control more difficult: Herbst (2000) suggests that Zaire is intrinsically prone to rebellion because its population lives around the edges of the country, in contrast, say, to Egypt. To test this we calculated a Gini coefficient of population dispersion. For the calculation of the Gini coefficient we used the population data per 400km cell. Analogous to the income Gini coefficient, the Gini coefficient of population dispersion will be high if the country's population is concentrated in a relatively small area of the country. In fact, the concentration of the population is slightly *lower* in peaceful countries (0.57) than in war countries (0.6). The remaining geographic variables were population, population density and the proportion of the population living in urban areas. For all series the data source is World Bank (2000). Population is likely to be correlated with conflict risk. If two identical areas, each with a conflict risk of p, are treated as a single area, the conflict risk for the single area rises to $2p - p^2$. Since p is small (0.07 at the mean), this effect alone would yield an elasticity of conflict risk with respect to population of slightly less than unity. If there are scale economies in defense, the elasticity would be reduced. Empirically, peaceful countries have on average less than half the population of conflict countries, they have higher population density, and are more urbanized.

In war, both finance and geography may be less important than cohesion. The government army has two advantages over a rebellion. It can spend many years building a sense of unity, whereas if a rebel force fails to achieve unity quickly it will presumably perish. Additionally, the government can use the powerful rhetoric of nationalism: with this imagined identity already occupied, a rebellion needs an alternative. The need for rapid cohesion constrains recruitment: rebellion cannot afford diversity. There is now evidence that ethnic and religious diversity within various types of organization tends to reduce the ability to cooperate (Easterly and Levine, 1997, Alesina *et al*, 1999, Collier, 2001). A rebellion needs cooperative behavior and so it will tend to recruit only within a single ethnic/religious group. The need for homogeneity has a startling implication: rebellions should be more difficult the more diverse is the society since the recruitment pool becomes more restricted. The most widely used measure of ethnic diversity is the index of ethno-linguistic fractionalization which measures the probability of two randomly drawn people being from different ethnic groups.[4] On average, the conflict societies are *more* fractionalized than the peaceful societies. We could find no measure of religious fractionalization, but we constructed one equivalent to that of ethnic fractionalization

[4] This ethnolinguistic fractionalization index was for example used by Mauro (1995) and Easterly and Levine (1997, 1998). We relied upon the original source, Atlas Narodov Mira (USSR, 1964), and would like to thank Tomila Lankina for translating the data entries.

using data from Barrett (1982).[5] Religious fractionalization is similar in conflict and peaceful societies. Potentially, ethnic and religious diversity interact. If ethnic and religious divisions are cross-cutting then social fractionalization is much greater than the sum of ethnic and religious fractionalization. We could find no data on how religious and ethnic divisions are related and so we allow for the possibility that they are cross-cutting, a concept we refer to as social fractionalization. If there were *e* equally sized ethnic groups and *r* equally sized religious groups, maximum potential social fractionalization would be measured simply by the product $e \cdot r$. However, since both the underlying indices of ethnic and religious fractionalization range on the scale 0-100, their product is zero if there is either religious or ethnic homogeneity whereas there is social homogeneity only if both indices are zero. We therefore measure social fractionalization as the sum of the product of the underlying indices and whichever is the higher of them.

Grievance-Rebellion

We now switch from a focus upon the constraints facing rebellion to a focus upon preferences. To what extent is the initiation of rebellion determined by differences in objective grievances? We consider three grievances: inter-group hatred, political exclusion and vengeance.

Inter-ethnic, or inter-religious hatreds are probably the most common explanation for civil conflict. A possible example might be Bosnia. Although such hatreds cannot be quantified, they can evidently only occur in societies which are multi-ethnic or multi-religious. We have already discussed measures of ethnic and religious fractionalization: inter-group hatreds must be greater in societies which are fractionalized than in those which are homogenous. However, such hatreds need not be monotonic in fractionalization because polarized societies may generate more tensions than highly fractionalized societies. Esteban and Ray (1999) and Reynal-Querol (2000) suggest that it is polarization which increases the risk of conflict. Esteban and Ray (1994) consider how polarization can be measured. They show that the allowable class of measures is quite limited. Using data on ethno-linguist groups (source: *Atlas Narodov Mira*, USSR, 1964) we adopt their general measure of polarization:

$$P = K \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i^{1+\alpha} \pi_j d$$

where π_i denotes the percentage of people that belong to group *i* in the total population, *i=1, …, n*. This measure of polarization depends on the parameters *K* and α. *K* does not change the order, but is used for population normalization. For $\alpha=0$ the polarization measure is equal to the Gini coefficient of inequality. Esteban and Ray (1994) show that α is bounded, otherwise not all of their axioms are fulfilled. They define the boundaries for α as : $0 < \alpha \leq \alpha^*$ where $\alpha^* \cong 1.6$. The degree of antagonism between two different ethnic groups is denoted by *d*. Obviously, in large samples such

[5] We would like to thank Robert Barro who made most of this data available in electronic format. We followed his suggestions to aggregate the various religious affiliations (Barro, 1997).

as we are using this is not observed. Following Reynal-Querol (2000) we assume that the distance between any two different ethnic groups is one whereas that within the group is zero, so that *d* has the properties:

d=1 if i≠j and *d=0 if i=j*.

We calculate the polarization measure for three different values of α : α = 0, α =0.8 and α =1.6. In addition we investigate the variant of the Esteban-Ray measure which Reynal-Querol proposes. The four polarization measures are highly correlated with each other, but indeed measure something distinct from fractionalization for $\alpha > 0$. For α = 0 the polarization measure is equivalent to the ethnic fractionalization index. Thus, the correlation coefficient between polarization and fractionalization ranges between 1.0 (α = 0) and 0.39 (α = 1.6). The descriptive statistics in Table 2 indicate that war and non-war countries had very similar values of polarization, the means and the standard deviations are virtually identical.

Political exclusion can occur either through generalized repression or because some particular group is victimized. We measure political repression using the Polity III data set (Jaggers and Gurr, 1995). Their democracy measure ranges 0-10 on an ascending scale of openness. There is a very large difference in the extent of democracy between conflict societies and peaceful societies: on average, conflict episodes are preceded by a democracy score less than half that which precedes peace episodes. We also investigate the Polity III measure of autocracy, and a measure of political openness published by Freedom House (the `Gastil Index'). The quantitative political science literature has already used these measures to explore the relationship between conflict and the political rights of a society. Gleditsch and Hegre (1997) find that other than when repression is very severe it tends to increase the risk of conflict.

However, even democracies may exclude if a minority is systematically victimized. One circumstance in which a stable winning coalition can form is if one ethnic group constitutes a majority. The incentive to exploit the minority diminishes the smaller is the minority, since there is less potential for exploitation. Hence, the most likely range for ethnic exclusion is if the largest ethnic group constitutes a majority, but not an overwhelming majority, of the population. We refer to this as ethnic dominance. Empirically, we specify some range for the largest group, eventually choosing 45-90 percent of the population, and set an ethnic dominance dummy variable equal to unity for societies so characterized. The descriptive statistics suggest that ethnic dominance is unimportant: the proportion of conflict societies characterized by ethnic dominance is the same as that of peaceful societies.

A second circumstance of political exclusion is if the poor are marginalized. As Hirschleifer (1991) shows, normally the poor will succeed in using the political contest to ameliorate their economic position. A high degree of economic inequality is therefore some indication that the poor are atypically marginalized. The opening page of Sen's book *On Economic Inequality* asserts that `the relation between inequality and rebellion is indeed a close one'. The `rage of the poor' is probably the most popular explanation of conflict after that of ethnic hatred, and may be exemplified by the Castro rebellion in Cuba. However, the case study evidence casts doubt on the hypothesis: a survey of fifteen violent civil conflicts concludes that `wars today are rarely started by poor and marginalized people united in battle as an expression of

their deep-seated striving for a just society.' (Anderson, 1999, p.9). An alternative route by which inequality might induce conflict is if rich regions pre-empt redistribution by attempting secession (Buchanan and Faith, 1987). Empirically, we investigate the effect of the inequality of both income and land. For income we investigate both the Gini coefficient and the ratio of the top-to-bottom quintiles of income, and for land we use the Gini coefficient. The data is from Deininger and Squire (1996, 1998). The conflict societies are on average slightly more unequal than the peaceful societies.

Finally, we consider grievance due to a history of previous conflict. We have already discussed measures of previous conflict. However, while the greed model has used these measures to proxy rebellion-specific capital, the grievance model would interpret the same variables as proxies for intensified hatred. Conflict evidently generates hatred, and hatred is likely to fade with time.

Greed-Grievance Comparisons

As set out above, the proxies for constraints and the proxies for preferences are largely distinct and so can be compared as two non-nested econometric models. There is, however, no reason for the accounts to be exclusive and the aim of our econometric tests is to arrive at an integrated model which gives an account of conflict risk in terms of all those constraints and preferences which are significant.

In interpreting the results a potential problem is if the same variable appears in both greed and grievance accounts. Ethnicity and religion enter both accounts, but the predictions are sufficiently different to be distinguishable. Entering as a constraint, fractionalization should make rebellion more difficult, and if ethnic and religious divisions are cross-cutting, the individual fractionalization measures should be dominated by our measure of social fractionalization. Entering via preferences, both ethnic and religious fractionalization should increase the risk of rebellion, and should be dominated by our measures of polarization. Ethnic dominance potentially proxies both ethnic hatred and objective political grievance. However, since our polarization measures should be better proxies for ethnic hatred, any remaining effect of ethnic dominance might reasonably be interpreted as being due to political exclusion. A more serious problem of interpretation is posed by the history of previous conflict. In the greed account this proxies rebellion-specific capital while in the grievance account it proxies hatred. We attempt to resolve this by testing for an interaction effect which could only be interpreted as rebellion-specific capital.

A second problem of interpretation is that some variables have so many possible effects that even if they are statistically significant, the particular route by which they effect conflict risk must remain speculative. This is particularly pertinent for per capita income and secondary school enrolment. We return to these issues after establishing the econometric model.

3. Results

Our empirical analysis attempts to predict the risk that a civil war will start during a five-year sub-period, through a logit regression in which the explanatory variables are characteristics at the start of the sub-period. Our data set describes a country in 1960, 1965…1995 by a vector of characteristics which appertained either in that year or during the five years. We consider only countries which were at peace in that year and predict whether the peace was sustained through to the end of 1964, 1969…1999, respectively. Our model results depend on how societies which experienced outbreaks of war differed from the societies which sustained peace.[6]

We start with the greed model (see Table 3). Because per capita income and enrolment in secondary schooling are highly correlated, they cannot be used in the same regression. The first four columns include secondary schooling but not per capita income, which permits a sample of 123 countries. There are up to seven observations per country and our sample consists of 688 episodes of which 46 are war observations. The fifth column replicates the core regression using per capita income instead of secondary schooling, which permits a sample of 750 episodes of which 52 are war observations. Recall that there have been 78 civil wars that meet our criteria: 26 wars must be excluded due to missing data.

The first column omits the effects of previous conflict. The two opportunity cost proxies are significant with the expected signs. A higher gross secondary school enrolment rate for males reduces the risk of war. Income growth per capita, measured for the previous five year period, decreases the risk of war. The effect of primary commodity exports on the risk of war is significant but non-linear. The highest risk of war is at a share in GDP of about 32 percent. The positive linear effect on risk is consistent with the hypothesis that ability to raise finance is a determinant of rebellion. The negative quadratic effect suggests that the increased tax revenue eventually augments the capacity of the government to defend itself sufficiently to offset the enhanced finances of the rebels. The other proxy for the ease of financing of rebellion, the dummy variable for the end of the Cold War, is negative as expected, but is insignificant. The three geography variables are all significant with expected signs. The elasticity of the risk of conflict with respect to population is less than unity, consistent with modest economies of scale in defense. More mountainous countries are more likely to experience a civil war, while a higher concentration of the population decreases the risk of conflict. Social fractionalization enters with the expected negative sign but is not significant.

[6] In our earlier work (Collier and Hoeffler, 1998) we used a much smaller sample of wars and conflated war starts with war duration, using a tobit procedure. We now regard this approach as flawed since duration seems to be determined by rather different factors from starts.

Table 3: Greed Model

	1	2	3	4	5	6	7
Male secondary schooling	-0.0312 (0.010)***	-0.029 (0.010)***	-0.025 (0.010)**	-0.024 (0.010)***			
Ln GDP per capita					-0.837 (0.253)***	-1.237 (0.283)***	-1.243 (0.284)***
GDP growth	-0.119 (0.044)***	-0.116 (0.043)***	-0.117 (0.044)***	-0.118 (0.044)***	-0.105 (0.042)***		
Primary commodity exports/GDP	19.990 (5.882)***	17.634 (5.959)***	18.149 (6.006)***	18.900 (5.948)***	16.476 (5.207)***	17.567 (6.744)***	17.404 (6.750)***
(Primary commodity exports/GDP) ^{2}	-31.562 (12.003)***	-26.171 (11.889)**	-27.445 (11.996)***	-29.123 (11.905)***	-23.017 (9.972)**	-28.815 (15.351)*	-28.456 (15.366)*
Social fractionalization	-0.0001 (0.0001)	-0.0002 (0.0001)*	-0.0002 (0.0001)	-0.0002 (0.0001)	-0.0002 (0.0001)**		
Previous war		1.057 (0.374)***	0.464 (0.547)				
Peace duration			-0.003 (0.002) p=0.128	-0.004 (0.001) ***	-0.004 (0.001)***	-0.002 (0.001)	-0.002 (0.001)
Post-coldwar	-0.518 (0.427)	-0.588 (0.434)	-0.326 (0.469)	-0.207 (0.450)	-0.454 (0.416)		
Diaspora/peace						700.931 (363.29)**	
Diaspora corrected/peace							741.168 (387.635)*
(Diaspora-diaspora corrected)/peace							82.798 (287.192)
Ln population	0.849 (0.155)***	0.710 (0.161)***	0.669 (0.163)***	0.686 (0.162)***	0.493 (0.129)***	0.295 (0.141)**	0.296 (0.141)**
Geographic dispersion	-2.281 (1.014)**	-2.394 (1.024)**	-2.211 (1.038)**	-2.129 (1.032)**	-0.865 (0.948)		
Mountainous terrain	0.016 (0.008)**	0.012 (0.009)	0.013 (0.009)	0.014 (0.009)	0.008 (0.008)		
N	688	688	688	688	750	595	595
No of wars	46	46	46	46	52	29	29
Pseudo R^2	0.21	0.23	0.24	0.24	0.22	0.25	0.25
Log likelihood	-133.79	-129.69	-128.49	-128.85	-146.86	-93.27	-93.23

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively.

In column 2 we introduce a dummy variable for previous conflicts. Recall that this may proxy `rebellion-specific capital' or simply reflect omitted country-specific effects. Here we include any civil wars which occurred between the end of World War II and 1995. The coefficient on this previous war dummy is highly significant: previous wars increase the risk of new civil war. In column 3 we distinguish between the two interpretations of the previous war dummy by adding to the model the length of the peace period measured in months since the end of the last civil war. The coefficient is negative and significant, i.e. the longer the peace period the lower the risk of conflict. The coefficient on the previous war dummy is now insignificant. Thus, the peace period is a more precise measure of the effect than is the previous war dummy (p=0.59). This indicates that the risk decays after conflict, as might rebel-specific capital, rather than being a proxy for unobserved country-specific effects. In column 4 we therefore drop the insignificant previous war and post-Cold War dummies, leaving a model in which greed causes initial conflict; conflict causes some effect such as the accumulation of rebellion-specific capital, and this causes further conflict. In the last column we present an alternative specification in which we include the average income per capita as a measure of opportunity costs instead of schooling. The results are very similar, although the overall performance of the regression slightly deteriorates. The addition of this previous war effect has one important consequence for the other variables in the greed model: the negative effect of social fractionalization on conflict risk now becomes statistically significant. Fractionalized societies are actually safer than homogenous societies.

Recall that the duration of previous peace may proxy either rebellion-specific capital or hatred. In Table 3 we attempt to distinguish between these interpretations by investigating whether the effect of the duration of previous peace is itself affected by the size of the diaspora. Since the data set on the size of diasporas in the USA reduces the number of countries on which we have data, the sample size is consequentially radically reduced from the 750 observations and 52 wars which is our maximum sample in the previous analysis. In order to preserve sample size we therefore retreat to a more parsimonious version of the model, dropping four sample-constraining peripheral explanatory variables: ethnic and religious fractionalization, geographic concentration of the population, the extent to which the terrain is mountainous, and the rate of growth in the previous five year period. The remaining explanatory variables are thus per capita GDP, primary commodity exports, population, and the number of months since the previous conflict. Even with these data-restoring deletions, the sample size is reduced to 595 observations (containing 29 war observations). However, all the included explanatory variables remain significant.

On this sample, we then test for the effect of a diaspora. We measure the size of the diaspora relative to the resident population in the country of origin. To allow for a fading post-conflict effect, we interact this measure with the number of months since the previous conflict, dividing the former by the latter. This variable, `diaspora/peace' in Table 3, is added to the regression, the result being shown in column 6. The variable is positive and significant, while the duration of peace itself is not significant at the conventional level (p=0.14). A large diaspora considerably increases the risk of further conflict. If we compare the post-conflict society with the largest diaspora against that with the smallest, with other variables at their mean values, after five years of peace the risk of renewed conflict is around six times greater.

However, while this result is consistent with the analytic model, it is also open to a more anodyne interpretation. Evidently, diasporas are to an extent endogenous to the intensity of conflict: when civil war occurs, many people leave and settle in the USA. Hence, the size of the diaspora might simply be proxying the intensity of conflict. The result may therefore simply show that more severe conflicts have a higher risk of renewed conflict. To test for this we decomposed observed diasporas into a component which is exogenous to the intensity of conflict and a residual component which is therefore endogenous to its intensity. For this decomposition we estimated a simple migration model, reported in Appendix A1. The size of the diaspora in a census year is predicted to be a function of its size in the previous census, time, per capita income in the country of origin, and whether there was a war in the intervening period. This model predicts the size of the diaspora with reasonable accuracy. We then replace the diaspora data used in the model with estimated diaspora size in all cases where the observed diaspora is for a year subsequent to a conflict. Thus, all post-conflict observations of diasporas are estimates which are purged of any effect from the intensity of conflict. The difference between these estimates and the actual figures are then used as an additional variable, measuring that part of the diaspora which is potentially endogenous to the intensity of conflict. Both of these measures are then introduced into the regression in place of the previous single measure of the diaspora. The results are reported in column 7 of Table 3. The purged measure of the diaspora remains significant, and the size of the coefficient is only slightly altered. Further, its coefficient is not significantly different from that on the endogenous diaspora measure. Had the effect of the diaspora been simply a proxy for the intensity of conflict, neither of these would have been the case. The purged variable would have been insignificant, and the coefficient on the endogenous measure would have been larger. This suggests that the substantial effect of the diaspora on the risk of conflict renewal is indeed due to its financial contribution to war start-up. This result is not only of interest for its perspective on diasporas, it also helps to distinguish between the `rebellion-specific capital' and hatred interpretations of the duration of peace effect. If the duration of peace effect is proxying rebel-specific capital, it is easy to see why there would be such a strong interaction effect from diasporas. A rebel organization can be sustained financially during a period of post-conflict peace by contributions from the diaspora: hence, rebel organizational capital decays less rapidly the larger is the diaspora. For the hatred interpretation to remain in contention we would need to believe that diasporas substantially slow the rate of decay of hatred among the non-diaspora population. While some such effect is possible, diasporas are small and distant. It is more plausible to interpret their disproportionate influence as being due to their genuinely disproportionate financial power, rather than to some assumed disproportionate influence on attitudes.

Table 4: Grievance Model

	1	2	3
Ethnic fractionalization	0.010 (0.006)*	0.011 (0.007)*	0.012 (0.008)
Religious fractionalization	-0.003 (0.007)	-0.006 (0.008)	-0.004 (0.009)
Polarization α =1.6	-3.067 (7.021)	-4.682 (8.267)	-6.536 (8.579)
Ethnic dominance (45-90%)	0.414 (0.496)	0.575 (0.586)	1.084 (0.629)*
Democracy	-0.109 (0.044)***	-0.083 (0.051)*	-0.121 (0.053)**
Peace duration	-0.004 (0.001)***	-0.003 (0.001)***	-0.004 (0.001)***
Income inequality		0.015 (0.018)	
Land inequality			0.461 (1.305)
Ln population	0.221 (0.096)**	0.246 (0.119)**	0.300 (1.133)**
Geographic dispersion	-0.509 (0.856)	-0.763 (1.053)	-1.293 (0.102)
Mountainous Terrain	0.011 (0.007)	0.007 (0.009)	-0.0001 (0.009)
N	850	604	603
No of wars	59	41	38
Pseudo R^2	0.13	0.11	0.17
Log likelihood	-185.57	-133.46	-117.12

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively.

Column 1: the two measures of fractionalization and ethnic dominance are not jointly significant.

In Table 4 we turn to the examination of a rebellion which is motivated only by grievance. In the first column we examine the relationship between ethnic dominance, ethnic and religious fractionalization, ethnic polarization, democracy and the duration of peace. At this stage we define ethnic dominance as occurring when the largest ethnic group constitutes 45-90 percent of the population and measure polarization with α = 1.6. These specifications are justified in Section 4 where we investigate robustness to alternative definitions. As in the greed model, we control for geographic military advantage by including population, the dispersion of the population, and mountainous terrain. Since we are not including any lagged variables we can use 850 observations of which 59 observations experienced an outbreak of civil war. The results suggest that a higher degree of ethnic fractionalization increases the risk of war and that a greater openness of political institutions reduces the risk of conflict. Religious fractionalization, ethnic polarization and ethnic dominance are neither

individually nor jointly significant. In the second column we add the Gini coefficient of income inequality as an explanatory variable and in column 3 we add the Gini coefficient of land inequality as an alternative measure of inequality. Neither measure is significant. Note that the sample size is reduced when we include the income inequality measures. However, we are still analyzing a substantial number of war occurrences (41 in column 2 and 38 in column 3).[7] All three grievance models have very low explanatory power, the regressions only have an R^2 of 0.17 or lower.

We now turn to the question of which model, greed or grievance, provides a better explanation of the risk of civil war. Since the two models are non-nested, i.e. one model is not a special case of the other, we use the J-test as suggested by Davidson and MacKinnon (1981). This non-nested test is based on the following artificial nesting procedure. First we explain the risk of civil war, *p*, in terms of the two different models, greed and grievance.

(1) p=f(greed)
(2) p=f(grievance)

Based on these logit regressions we calculate the predicted probabilities and add these predicted values, $\hat{p}^{\text{greed}}$ and $\hat{p}^{\text{grievance}}$ to our alternative models.

(1) p=f(greed, $\hat{p}^{\text{grievance}}$)
(2) p=f(grievance, $\hat{p}^{\text{greed}}$)

According to the J-test the significance of the coefficients of these added variables enables us to choose between the two different models. If $\hat{p}^{\text{grievance}}$ is significant in the greed model we reject the greed model in favor of the grievance model. If $\hat{p}^{\text{greed}}$ is significant in the grievance model we reject the grievance model in favor of the greed model. As can be seen in columns 1 and 2 of Table 5, $\hat{p}^{\text{grievance}}$ is significant in the greed model and $\hat{p}^{\text{greed}}$ is significant in the grievance model. Thus, we conclude that while the greed model is superior, some elements of the grievance model are likely to add to its explanatory power. We therefore investigate the combination of the two models as presented in column 3 of Table 5.

Since this combined model includes income inequality and a lagged term our sample size is much reduced (479 observations). Omitting inequality (which remains insignificant) increases the sample size to 665 observations in column 4. In this combined model neither democracy, ethnic and religious fractionalization nor the post-Cold War dummy are significant. Other variables are statistically significant or close to significance and the overall fit is reasonable (pseudo R^2 of about 0.26). Since both the grievance and greed models are nested in the combined model, we can use a likelihood ratio test to determine whether the combined model is superior. We can reject the validity of the restrictions proposed by the grievance model, but not by the greed model.[8]

[7] We also tried the ratio of the income shares of the top to the bottom quintiles. This was also insignificant.

[8] Using the same sample as for the combined model (n=665) we obtain the following results: Greed model versus combined model, 5 degrees of freedom, Likelihood Ratio Test (LRT): χ^2 =7.85, p=0.165; grievance model versus combined model, 6 degrees of freedom, LRT: χ^2 =29.64, p= 0.000.

Table 5: Combined Greed and Grievance Model

	1	2	3	4	5
Male secondary schooling	-0.021 (0.011)**		-0.029 (0.013)**	-0.022 (0.011)**	-0.023 (0.011)**
Ln GDP per capita					
(GDP growth)t-1	-0.108 (0.044)***		-0.045 (0.062)	-0.108 (0.045)**	-0.103 (0.044)**
Primary commodity exports/GDP	19.096 (5.993)***		37.072 (10.293)***	23.385 (6.692)***	23.204 (6.660)***
(Primary commodity exports/GDP)^2	-30.423 (12.008)***		-69.267 (21.697)***	-36.335 (12.998)***	-36.206 (12.946)***
Social fractionalization	-0.0002 (0.0001)***		-0.0008 (0.0003)**	-0.0005 (0.0003)	-0.0005 (0.0003)
Ethnic fractionalization		0.008 (0.007)	0.041 (0.019)**	0.023 (0.015)	0.022 (0.015)
Religious fractionalization		-0.005 (0.008)	0.015 (0.020)	0.014 (0.019)	0.014 (0.019)
Polarization α =1.6		-9.358 (8.735)	-25.276 (13.390)*	-15.992 (10.518)	-15.556 (10.476)
Ethnic dominance (45-90%)		1.212 (0.648)**	2.020 (0.915)**	1.592 (0.746)**	1.556 (0.740)**
Democracy		-0.036 (0.054)	-0.018 (0.062)	-0.042 (0.054)	-0.044 (0.054)
Peace duration	-0.0003 (0.002)	0.0005 (0.0014)	-0.0003 (0.0015)	-0.003 (0.001)***	-0.003 (0.001)***
Post-coldwar	-0.209 (0.457)		-0.873 (0.644)	-0.281 (0.459)	
Income inequality			0.025 (0.024)		
Ln population		-0.014 (0.136)	0.927 (0.250)***	0.697 (0.181)***	0.685 (0.179)***
Geographic dispersion	-1.978 (1.049)*	0.135 (1.106)	-4.032 (1.490)***	-1.962 (1.149)*	-1.957 (1.153)*
Mountainous Terrain	0.005 (0.010)	0.001 (0.008)	0.005 (0.012)	0.015 (0.009)	0.014 (0.009)
Grievance predicted value	0.767 (0.413)**				
Greed predicted value		1.052 (0.212)***			
N	665	665	479	665	665
No of wars	46	46	32	46	46
Pseudo R^2	0.24	0.25	0.24	0.26	0.25
Log likelihood	-126.69	-125.29	-89.55	-124.60	-124.79

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively

Although the combined model is superior to the greed and grievance models, several variables are completely insignificant and we drop them sequentially. First we exclude the post-Cold War dummy, then religious fractionalization, then democracy[9], then polarization, then ethnic fractionalization (column 9). Social fractionalization and mountains are both marginally significant in this model (p-value around 0.13) and are jointly significant. When either is dropped, the other becomes significant and in the present model there is little to choose between them. However, when we switch to the larger sample permitted by replacing male secondary school enrolment with per capita income, there is a clear ranking. When both variables are included, social

[9] We tried different specifications to test for the effect of political repression by investigating non-linear effects, by including the autocracy score instead of the democracy score, and by using the difference between the two variables as suggested by Londregan and Poole (1996). We also tried the Freedom House measure of political freedom, but neither of these alternative political repression measures were found to be significant.

Table 5 continued

	6	7	8	9	10	11
Male secondary schooling	-0.023 (0.011)**	-0.025 (0.011)**	-0.027 (0.010)***	-0.027 (0.010)***	-0.031 (0.010)***	
Ln GDP per capita						-0.950 (0.245)***
(GDP growth)t-1	-0.104 (0.044)**	-0.109 (0.044)***	-0.118 (0.044)***	-0.119 (0.043)***	-0.115 (0.043)***	-0.098 (0.042)**
Primary commodity exports/GDP	22.026 (6.385)***	22.439 (6.422)***	20.275 (6.126)***	20.107 (6.040)***	18.937 (5.865)***	16.773 (5.206)***
(Primary commodity exports/GDP)^2	-34.502 (12.617)***	-35.010 (12.760)***	-31.885 (12.358)***	-30.959 (12.076)***	-29.443 (11.781)***	-23.800 (10.040)**
Social fractionalization	-0.0003 (0.0001)**	-0.0003 (0.0001)**	-0.0002 (0.0001)*	-0.0001 (0.0001)	-0.0002 (0.0001)**	-0.0002 (0.0001)***
Ethnic fractionalization	0.015 (0.010)	0.015 (0.010)	0.009 (0.009)			
Polarization α =1.6	-12.818 (9.779)	-13.192 (9.738)				
Ethnic dominance (45-90%)	1.373 (0.688)**	1.379 (0.692)**	0.602 (0.365)*	0.683 (0.356)*	0.670 (0.354)*	0.480 (0.328) p=0.14
Democracy	-0.042 (0.054)					
Peace duration	-0.003 (0.001)***	-0.003 (0.001)***	-0.003 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***
Ln population	0.684 (0.179)***	0.699 (0.177)***	0.724 (0.172)***	0.749 (0.169)***	0.768 (0.166)***	0.510 (0.128)***
Geographic dispersion	-2.211 (1.099)**	-2.097 (1.078)**	-2.379 (1.050)**	-2.224 (1.029)**	-2.487 (1.005)**	-0.992 (0.909)
Mountainous Terrain	0.015 (0.009)	0.015 (0.009)*	0.013 (0.009)	0.014 (0.009)		0.014 (0.009)
N	665	688	688	688	688	750
No of wars	46	46	46	46	46	52
Pseudo R^2	0.25	0.26	0.25	0.25	0.24	0.22
Log likelihood	-125.05	-125.73	-126.67	-127.11	-128.21	-146.84

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively

fractionalization is highly significant, whereas mountains are insignificant. We therefore drop mountains yielding the baseline model of column 10 and its variant with per capita income replacing secondary enrolment in column 11. No further reduction in the model is accepted, and no additions of variables included in our previous models are accepted. The baseline model and its variant yield very similar results although the variant has less explanatory power and two variables lose significance (ethnic dominance and geographic dispersion). That male secondary school enrolment and per capita income yield such similar results suggests that several interpretations may be consistent with the results. While male enrolment may proxy the opportunity cost of joining a rebel movement, it may also proxy other effects of income and education.

Our baseline model allows us to calculate the change in the probability of war-starts for different values of the explanatory variables. We present these calculations in Appendix Table A2. At the mean of all variables the risk of a war-start is about 11.5 percent. Our model predicts that a country with the worst characteristics (lowest per capita income, lowest GDP growth and highest population growth, a primary commodity export share of 0.32, the largest population, the lowest fractionalization, ethnic dominance, a geography Gini coefficient of zero and only one month of peace) would have a near-certain risk of war while a country with the best characteristics

would be a very safe society. We now calculate how the change in one variable (while the others are assumed to take their mean values) affects the probability of experiencing a civil war.

The two opportunity cost proxies, economic growth and male secondary school enrollment, both have substantial effects. An additional percentage point on the growth rate reduces the risk of war by around two percentage points (a decline from 11.5 percent to 9.3 percent). If the enrollment rate is 10 percentage points higher than the average the risk of war is reduced by about three percentage points (a decline in the risk from 11.5 percent to 8.6 percent). Alternatively, enrolment may proxy the level of economic development. A doubling of per capita income reduces the risk of war by five percentage points. The effect of primary commodity exports is considerable: at the risk-maximizing value of the primary commodity export share (0.32), the risk of civil war is about 22 percent, while a country with no natural resource exports only has a probability of a war-start of one percent. Hence, the basic economic variables - the level, growth and structure of income - are highly important in conflict risk.

Consistent with the hypothesis that cohesion is important for rebel effectiveness, social fractionalization makes a society substantially safer: a maximally fractionalized society has a conflict risk only one quarter that of a homogenous society. However, if a country is characterized by ethnic dominance its risk of conflict is nearly doubled. Thus, the net effect of increased social fractionalization need not be monotonic. Starting from homogeneity, as fractionalization is increased a society is likely to become characterized by ethnic dominance, although this will be removed by further fractionalization. Directly after a civil war there is a high probability of a re-start, the risk is about 32 percent. This risk declines over time at around one percentage point per annum. Both ethnic dominance and the length of time since a previous conflict could be proxies for hatred. However, two results suggest that other interpretations are more plausible. First, neither polarization nor ethnic and religious fractionalization are significant risk factors. Thus, ethnic dominance is perhaps more likely to be proxying structural political exclusion than inter-group hatreds. Secondly, recall that the rate at which conflict risk declines during post-conflict peace is highly sensitive to the size of the diaspora. This suggests that during periods of conflict societies accumulate rebellion-specific capital and organizations.

Geography matters. The elasticity of the risk of conflict with respect to population is less than unity (significantly so in the income variant, though not quite in the baseline). This is consistent with there being economies of scale in deterrence. Lastly, countries with a highly concentrated population have a very low risk of conflict, whereas those with a highly dispersed population have a very high risk of civil war (about 37 percent).

4. Robustness Checks

We now test these baseline results for robustness. We consider the sensitivity both to data and to method. With respect to data, we investigate the effect of outlying observations, and of different definitions of the dependent and independent variables. With respect to method, we investigate random effects, fixed effects and rare events bias.

We investigated outlying observations through two approaches, graphs of the 46 conflict episodes used in the baseline regression and an analysis of influential data points. Because our sample is unbalanced as between events and non-events, the potential problems of outliers arises predominantly among the 46 conflict episodes used in the baseline regression rather than the 642 peace episodes. Further, of these conflict episodes, 24 were first-time conflicts and 22 were repeat conflicts. We first investigate these conflict episodes graphically, the first-time conflicts being shown by dots and the repeat conflicts by triangles. The mean of the peace episodes is shown by a cross.

In Figure 1 we plot the level and structure of income: that is, per capita income against the primary commodity export ratio. Two observations, Iran in 1970 and 1975, are outliers in both dimensions. Especially since these two observations are for the same country at consecutive periods, results that were dependent upon these observations could not be regarded as general. Two other observations, Angola in 1990 and Congo in 1995, have high ratios of primary commodity dependence relative to the other observations in the war sample. However, these observations are not outliers relative to the peace sample: 26 peace observations have higher primary commodity dependence than the most extreme conflict observation.

In Figure 2 we plot male secondary school enrolment against the growth rate of per capita income, both being proxies for the opportunity cost of joining a rebel organization. One observation, Romania in 1989, is an outlier in both dimensions. According to the data, Romania during the late 1980s offered atypically attractive opportunities to young men in schooling and employment. We may well be suspicious of the Romanian data and so results that were dependent upon this observation would themselves be highly doubtful.

Thus, Iran and Romania appear to be doubtful observations and so we re-estimate the baseline regressions dropping these countries (Table 6, column 1). No results are overturned, but the performance of the regression improves and variables are more significant. While this tends to support the diagnosis of these observations as unreliable, the more important conclusion is that these outliers do not account for the results.

Figure 1: Income Structure and Level

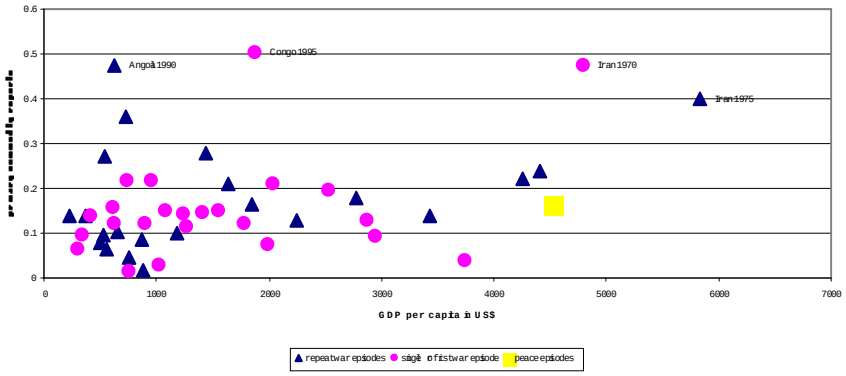


Figure 2: Male Secondary Schooling and Income Growth

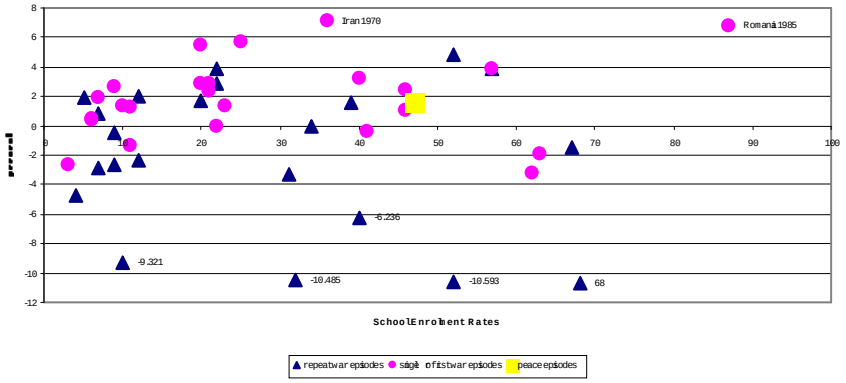


Figure 3: Religious and Ethnic Fractionalization

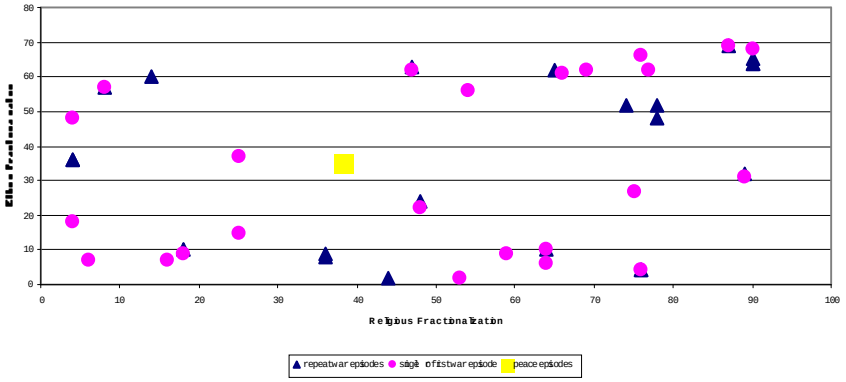


Figure 4: Geographic Dispersion and Mountainous Terrain

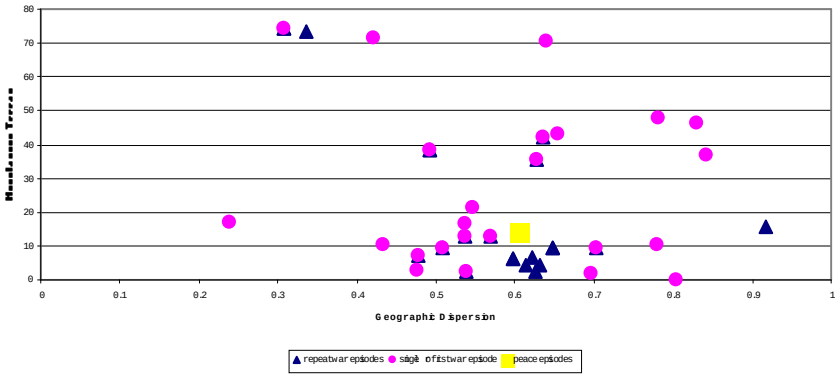


Figure 5: Democracy and Income Distribution

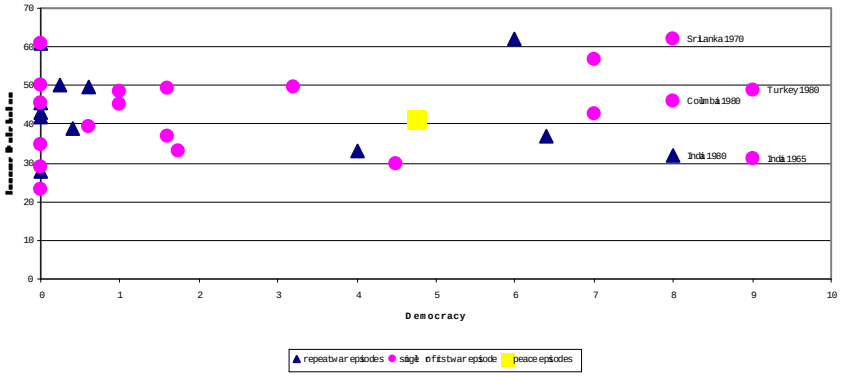
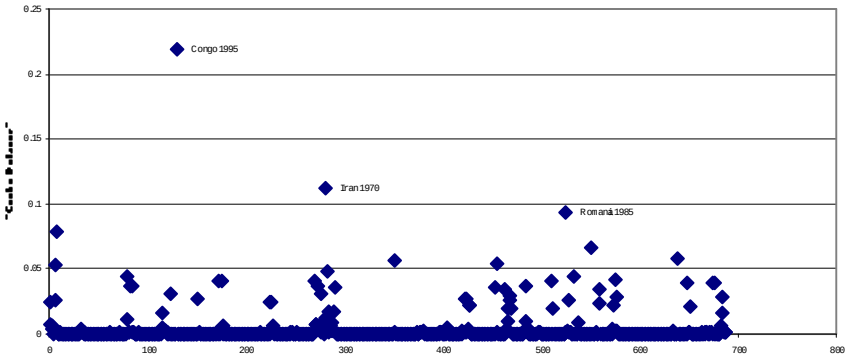


Figure 6: Influential Data Points



There are four observations of highly negative growth: Angola in 1970-74, Zaire in 1990-95, Iran in 1975-79 and Iraq in 1980-84. All of these growth collapses appear to be genuine, and they occur in different countries. However, were the results to be dependent upon these four observations, an implication would be that the risk of conflict was not monotonic in growth, but rather influenced by severe collapses. All four of these growth collapses are cases in which the subsequent conflict was a repeat. Were first-time conflict to be collinear with subsequent growth collapses, a more general problem of interpretation would arise. However, both analytically and empirically, post-conflict situations can be characterized by atypically high growth as well as by decline (Collier, 1999). We investigate whether the four extreme growth collapses account for the effect of growth on conflict risk, deleting them along with Iran and Romania (Table 6, column 2). Growth remains significant, and its coefficient is only slightly reduced. Hence, we can conclude that the increased risk of conflict due to slow growth is not confined to episodes of growth collapse, but is a more continuous relationship.

In Figure 3 we plot religious and ethnic fractionalization. Although there is a wide dispersion of both measures, there are no obvious outliers. Figure 4 plots geographic dispersion against mountainous terrain. Again, there are no obvious outliers. In Figure 5 we plot two measures of grievance: democracy and income distribution. Again, although there is a wide dispersion, there are no obvious outliers, nor are the extreme conflict observations outliers relative to the peace observations. Nor are there obvious outliers in respect of diasporas. The conflict episode with the largest diaspora, Nicaragua in 1980, had 1.6 emigrants in the USA per 100 inhabitants. Since the mean for the peace observations was 1.4, we conclude that outliers cannot account for our result on the effect of diasporas on conflict risk.

Four peaceful countries have particularly high values of primary dependence: Saudi Arabia, Guyana, Oman, and Trinidad and Tobago. It is possible that these observations account for the non-monotonic relationship to conflict risk. This might imply that the reduction in conflict risk only occurred at extreme values of commodity dependence. In Table 6 column 3 we present our baseline model excluding the four high primary commodity exporters. The main results are unchanged and the non-monotonic relationship between primary commodity exports and the risk of conflict is confirmed.

Table 6: Investigation of Outliers

	1	2	3	4
	Excluding Iran and Romania	Excluding Iran and Romania and growth collapses ♠	Excluding high primary commodity exporters ♦	Excluding influential data points ♣
Male secondary schooling	-0.035 (0.011)***	-0.037 (0.011)***	-0.031 (0.010)***	-0.041 (0.011)***
(GDP growth) t-1	-0.140 (0.047)***	-0.100 (0.052)**	-0.122 (0.044)***	-0.137 (0.046)***
Primary commodity exports/GDP	19.696 (6.608)***	19.029 (6.671)***	18.771 (6.063)***	28.745 (7.862)***
(Primary commodity exports/GDP)^2	-34.090 (14.356)**	-33.250 (14.609)**	-28.466 (12.299)**	-59.818 (17.781)***
Social fractionalization	-0.0002 (0.0001)**	-0.0002 (0.0001)**	-0.0002 (0.0001)**	-0.0003 (0.0001)***
Ethnic Dominance	0.727 (0.368)**	0.732 (0.370)**	0.647 (0.354)*	0.655 (0.372)*
Peace duration	-0.004 (0.001)***	-0.003 (0.001)***	-0.004 (0.001)***	-0.004 (0.0011)
Ln population	0.747 (0.174)***	0.743 (0.175)***	0.772 (0.168)***	0.899 (0.195)***
Geographic dispersion	-2.114 (1.080)**	-2.272 (1.090)**	-2.449 (1.008)**	-2.890 (1.136)***
N	674	671	662	685
No. of wars	42	39	46	43
Pseudo R^2	0.25	0.22	0.21	0.29
Log likelihood	-118.40	-116.17	-122.23	-114.04

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively

♠ We exclude the following three growth collapses: Angola 1970-74, Iraq 1980-84 and Zaire 1990-94.

♣ We exclude the following three influential data points: Iran 1970-74, Romania 1985-89, Congo 1995-99.

♦ We exclude the countries with the highest primary commodity export to GDP ratio, namely Saudi Arabia, Guyana, Oman and Trinidad and Tobago. Their average primary commodity export to GDP ratio is 0.504 (sample average 0.158).

We next analyze whether our regression results are sensitive to the inclusion of influential data points. Based on the methods developed by Pregibon (1981) we examined which observations may be influential and investigated whether omitting these observations from our baseline model changed our results. We used the following method to identify outliers.[10] First we predicted the probability of a civil war occurring, $\hat{p}_i$. Since the residuals of the regression, $y_i - \hat{p}_i$ are heteroskedastic we calculated the Pearson residual:

$$r_i = (y_i - \hat{p}_i) / \sqrt{\hat{p}_i(1 - \hat{p}_i)}$$

[10] Long (1997) pp.98-101 provides a discussion of influence in limited dependent variable models.

Large values of r_i indicate a poor fit for the *ith* observation. We also examined the standardized Pearson residual

$$r_i^{Std} = r_i / \sqrt{(1 - h_{ii})}$$

While these measures indicate outliers they do not indicate whether an observation has a large influence on the estimated coefficients or on the overall fit. On the other hand some observations can influence the estimates even if the residual is small. We searched for influential data points by computing an influence measure which is the counterpart to Cook's distance for the linear regression model

$$C_i = (\Delta_i \hat{\beta})' \hat{Var}(\hat{\beta})(\Delta_i \hat{\beta}) = r_i^2 h_{ii} / (1 - h_{ii})^2$$

where h_{ii} is a vector containing the diagonal elements of matrix H

$$h_{ii} = \hat{p}_i (1 - \hat{p}_i) x_i \hat{Var}(\hat{\beta}) x_i'$$

$$H = \hat{V} X (X' \hat{V} X)^{-1} X' \hat{V}$$

and where $\hat{V}$ is a diagonal matrix with $\sqrt{\hat{p}_i (1 - \hat{p}_i)}$ on the diagonal and $\hat{Var}(\hat{\beta})$ is the estimated covariance of the maximum likelihood estimator $\hat{\beta}$.

In figure 6 we plot this measure of influence. We find three influential observations: Congo 1997, Iran 1974 and Romania 1989. However, when we omitted these three observations from our regression, the overall fit of the regressions improved (from previously R2=0.24 to R2=0.29) and all of the coefficients remain statistically significant. (Table 6, Column 4).

We next turn to questions of variable definition. The most contentious aspect of the dependent variable is distinguishing between whether a country has a single long war or multiple shorter wars interrupted by periods of peace. In the above analysis we have been guided by the judgement of the political scientists who built the original data sets. Some peace periods are, however, quite short and it might be better to conceptualize these as interludes in a single war. We therefore reclassified all those wars which were separated by peace periods of less than one month, and then of less than one year, as continuous wars (Table 7A, column 1). The baseline results are not altered by the former redefinition, but when peace periods of less than twelve months are ignored (column 2) the effect of economic growth on conflict risk becomes

Table 7A: Variable Definitions: Wars and Social Fractionalization

	1	2	3	4	5	6	7	8	9	10
	Peace periods of shorter than one month are treated as continuous wars	Peace periods of shorter than one year are treated as continuous wars	Baseline Model plus ethnic fractionalization	ethnic fractionalization only	Baseline Model plus religious fractionalization	religious fractionalization only	Ethnic and religious fractionalization	Baseline Model plus ethnic and religious fractionalization	social frac.= ethnic plus religious fractionalization	social frac.= ethnic times religious fractionalization
Male secondary schooling	-0.031 (0.010)***	-0.031 (0.010)***	0.032 (0.010)***	-0.025 (0.009)***	-0.032 (0.010)***	-0.030 (0.010)***	-0.031 (0.010)***	-0.031 (0.010)***	-0.030 (0.010)***	-0.032 (0.010)***
(GDP growth) t-1	-0.102 (0.044)***	-0.071 (0.047)	-0.114 (0.043)***	-0.119 (0.043)***	-0.115 (0.043)***	-0.113 (0.043)***	-0.114 (0.043)***	-0.114 (0.043)***	-0.115 (0.043)***	-0.115 (0.043)***
Primary commodity exports/GDP	19.147 (5.939)***	22.686 (6.718)***	19.153 (5.970)***	15.283 (5.253)***	18.726 (5.894)***	17.082 (5.579)***	17.403 (5.643)***	19.426 (6.053)***	17.344 (5. 559)***	18.967 (5.871)***
(Primary commodity exports/GDP)^2	-30.150 (12.031)***	-39.053 (14.405)***	-30.420 (12.101)***	-22.833 (10.328)**	-29.330 (11.819)***	-27.020 (11.385)**	-27.361 (11.415)**	-30.810 (12.198)***	-26.623 (11.145)**	-29.499 (11.793)***
Social fractionalization	-0.0002 (0.0001)**	-0.0002 (0.0001)***	-0.0003 (0.001)**		-0.0002 (0.0002)			-0.0004 (0.0003)	-0.008 (0.004)	-0.0002 (0.0001)**
Ethnic fractionalization			0.009 (0.009)	-0.008 (0.006)			-0.003 (0.007)	0.011 (0.124)		
Religious fractionalization					-0.005 (0.013)	-0.017 (0.008)**	-0.016 (0.008)**	0.005 (0.017)		
Ethnic Dominance (45-90%)	0.732 (0.357)**	0.741 (0.362)**	0.594 (0.361)*	0.711 (0.352)**	0.646 (0.358)*	0.627 (0.357)*	0.645 (0.360)*	0.599 (0.361)*	0.727 (0.353)**	0.670 (0.354)*
Peace duration	-0.003 (0.001)***	-0.003 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***
Ln population	0.782 (0.167)***	0.832 (0.176)***	0.742 (0.169)***	0.672 (0.156)***	0.761 (0.167)***	0.709 (0.157)***	0.729 (0.166)***	0.743 (0.169)***	0.748 (0.164)***	0.768 (0.166)***
Geographic dispersion	-2.541 (1.012)***	-2.953 (1.049)***	-2.637 (1.024)***	-1.772 (0.948)**	-2.616 (1.057)***	-2.695 (1.049)***	-2.687 (1.050)***	-2.548 (1.063)**	-2.401 (1.005)**	-2.488 (1.005)***
N	686	683	688	688	688	688	688	688	688	688
No of wars	44	41	46	46	46	46	46	46	46	46
Pseudo R^2	0.23	0.21	0.24	0.22	0.24	0.24	0.24	0.24	0.23	0.24
Log likelihood	-126.33	-122.23	-127.75	-131.57	-128.13	-128.71	-128.64	-127.71	-129.17	-128.19

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively.
Column 1: We exclude the following war starts: Angola 1975 and Somalia 1988.
Column 2: We exclude the following war starts: Angola 1975, Mozambique 1976, Sierra Leone 1997, Somalia 1988 and Zaire/Democratic Rep. of Congo 1997.
Column 7: Ethnic and religious fractionalization are jointly significant at the ten percent level (p=0.091).

Table 7B: Variable Definitions: Ethnic Polarization and Ethnic Dominance

	1	2	3	4	5	6	7	8	9	10	11	12
	Number of Ethnic Groups	Baseline plus ethnic polarization	Ethnic polarization replaces social fractionalization	Ethnic polarization replaces ethnic dominance	Ethnic polarization replaces social fractionalization and ethnic dominance	Largest group as a proportion of the total population	45-85%	45-90%	45-95%	30-90%	40-90%	50-90%
Male secondary schooling	-0.029 (0.010)***	-0.031 (0.010)***	-0.023 (0.009)***	-0.031 (0.010)***	-0.023 (0.009)***	-0.029 (0.010)***	-0.029 (0.010)***	-0.032 (0.010)***	-0.029 (0.010)***	-0.032 (0.010)***	-0.031 (0.010)***	-0.031 (0.010)***
(GDP growth) t-1	-0.110 (0.043)***	-0.112 (0.432)***	-0.108 (0.043)***	-0.117 (0.043)***	-0.113 (0.043)***	-0.110 (0.043)***	-0.114 (0.043)***	-0.115 (0.043)***	-0.112 (0.043)***	-0.114 (0.043)***	-0.117 (0.043)***	-0.114 (0.043)***
Primary commodity exports/GDP	18.022 (5.829)***	19.660 (5.965)***	14.953 (5.198)***	17.745 (5.850)***	12.887 (5.076)***	18.258 (5.945)***	18.566 (5.849)***	18.937 (5.865)***	18.203 (5.783)***	18.812 (5.990)***	18.010 (5.870)***	18.582 (5.832)***
(primary commodity exports/GDP)^2	-28.137 (11.695)***	-30.307 (11.829)***	-22.594 (10.311)**	-27.920 (11.768)**	-20.076 (10.296)**	-29.136 (12.028)**	-28.982 (11.742)***	-29.443 (11.782)***	-28.065 (11.547)***	-30.159 (12.155)***	-28.232 (11.822)**	-29.105 (11.744)***
Social fractionalization	-0.0002 (0.0001)**	-0.0002 (0.0001)**		-0.0002 (0.0001)**		-0.0003 (0.0001)***	-0.0002 (0.0001)***	-0.0002 (0.0001)**	-0.0002 (0.0001)**	-0.0002 (0.0001)***	-0.0002 (0.0001)***	-0.0002 (0.0001)**
Ethnic Dominance		1.024 (0.660)	1.058 (0.650)*			-1.096 (1.015)	0.526 (0.355)	0.670 (0.354)*	0.264 (0.507)	0.598 (0.366)*	0.476 (0.349)	0.534 (0.361)
No of ethnic groups	-0.003 (0.013)											
Polarization α=1.6		-5.544 (8.615)	-4.685 (8.533)	5.704 (4.708)	6.903♠ (4.772) p=0.148							
Peace duration	-0.004 (0.001)	0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***	-0.004 (0.001)***
Ln population	0.728 (0.196)***	0.767 (0.167)***	0.636 (0.149)***	0.741 (0.163)***	0.615 (0.146)***	0.680 (0.161)***	0.740 (0.163)***	0.768 (0.166)***	0.715 (0.161)***	0.772 (0.167)***	0.736 (0.162)***	0.744 (0.163)***
Geographic dispersion	-2.390 (1.011)	-2.356 (1.026)**	-1.692 (0.968)*	-2.608 (1.014)***	-1.976 (0.959)**	-2.581 (1.025)***	-2.634 (1.024)***	-2.487 (1.005)***	-2.373 (1.007)**	-2.743 (1.022)***	-2.517 (1.005)***	-2.383 (0.998)**
N	688	688	688	688	688	688	688	688	688	688	688	688
No of wars	46	46	46	46	46	46	46	46	46	46	46	46
Pseudo R^2	0.23	0.24	0.22	0.23	0.22	0.23	0.24	0.24	0.23	0.24	0.24	0.24
Log likelihood	-129.98	-128.00	-130.97	-129.27	-132.37	-129.45	-128.94	-128.21	-129.88	-128.64	-129.07	-128.94

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively. We constructed the polarization measure following Esteban and Ray (1994). We assumed α=1.6. Our analysis showed that any other values of α produced results which were even less supportive of the hypothesis that polarization increases the risk of conflict.
Column 1: We never found the number of ethnic groups to be significant. We tried adding it to the baseline model, instead of the ethnic dominance variable (as shown above), instead of the social fractionalization variable and instead of both ethnic measures (social fractionalization and ethnic dominance).

marginally insignificant (p=0.12), although the coefficient is little effected. [11]

In the baseline we define ethnic dominance as the largest ethnic group constituting 45-90 percent of the population. In Table 7B we investigate six other definitions which either vary the range of the population or use the share of the largest group regardless of its size. As the range is changed from 45-90 percent the significance level and the coefficient are both reduced, while if the definition is changed more radically to being the population share of the largest group it is completely insignificant.[12] The range 45-90 percent is consistent with the theoretical prediction of the range over which a group which would have both the power and the incentive to exploit others (see Collier, 2001). In Table 7A we also demonstrate that `social fractionalization', our measure of cross-cutting cleavages, dominates the other possible aggregation procedures for ethnic and religious diversity. When this measure of fractionalization is included with the ethnic and religious diversity indices either together or individually, it is significant whereas the underlying indices are not significant.

In the baseline we use only the most extreme measure of polarization over the range proposed by Esteban and Ray (1994). However, if this measure is replaced by either the lower bound ($\alpha = 0$), or the central measure ($\alpha = 0.8$) the results are unaffected: polarization remains insignificant and the other variables remain significant. We also experimented with the alternative measure proposed by Reynal-Querol (2000), and with the number of ethnic groups, but with the same result.

In Table 8 we investigate a number of different estimation issues. We concentrate on the analysis of random effects, fixed effects, time effects and a correction for rare events. We re-estimated our models using random effects. For the baseline model we find that the panel data estimator is not different from the pooled estimator, i.e. we accept the hypothesis that we can pool across the observations.[13] The estimation of fixed effects logits was only possible on a very small sub-sample of the observations. The countries for which the dependent variable does not vary over time (the majority of countries experienced only peace) cannot be included in the analysis. Although the fixed effects test is very severe, the non-monotonic effect of primary commodity exports remains significant. Were the effect of primary commodity exports dependent only upon cross-section data, it might suggest that the variable was proxying some other characteristic such as geography. However, the fixed effects regression uses only changes in primary commodity dependence, and so reduces the scope for alternative interpretations.

[11] We also examined the effect of time since the previous conflict in more detail by including the natural logarithm of the peace variable or its square, however, a linear decay term provides a better fit. Note that the measure of peace since the end of the civil war is somewhat imprecise since we only measure it from the end of the war to the initial year of each sub-period. A duration model of post-war peace would allow a more detailed analysis of this peace effect, however, the duration model results in Collier, Hoeffler and Söderbom (2001) support the results presented in this paper.

[12] We also experimented with various other definitions of ethnic dominance, such as the dummy taking a value of one if the largest ethnic group made up 50-75, 50-80, 50-85 percent of the population. Neither of these definitions provided a better fit than our preferred definition based on 45-90 percent.

[13] A LRT provides a $\chi^2 = 0$ ($p = 0.998$). Thus, we cannot reject the null-hypothesis that the panel data and pooled estimator provide the same results.

Table 8: Estimation Issues

	1	2	3	4
	Random Effects	Fixed Effects	Pooled Logit plus Time dummies	Rare Events Logit
Male secondary schooling	-0.032 (0.010)***	0.007 (0.033)	-0.031 (0.010)***	-0.029 (0.010)***
(GDP growth) t-1	-0.115 (0.043)***	-0.045 (0.072)	-0.129 (0.047)***	-0.110 (0.040)***
Primary commodity exports/GDP	18.937 (5.865)***	35.850 (14.436)***	18.895 (5.988)***	17.161 (6.535)***
(Primary commodity exports/GDP)^2	-29.443 (11.782)***	-65.967 (26.964)***	-29.815 (12.098)***	-25.594 (14.355)*
Social fractionalization	-0.0002 (0.0001)**	-0.007 (0.006)	-0.0002 (0.0001)**	-0.0002 (0.0001)**
Ethnic Dominance (45-90%)	0.670 (0.354)*		0.682 (0.359)*	0.644 (0.336)*
Peace duration	-0.004 (0.001)***	0.011 (0.002)***	-0.004 (0.001)***	-0.004 (0.001)***
Ln population	0.768 (0.166)***	0.010 (1.410)	0.762 (0.170)***	0.726 (0.151)***
Geographic dispersion	-2.487 (1.005)***	115.363 (74.562)	-2.447 (1.018)**	-2.394 (1.085)**
T70-74			0.725 (0.602)	
T75-79			0.578 (0.608)	
T80-84			1.137 (0.602)*	
T85-89			-0.013 (0.757)	
T90-94			0.802 (0.677)	
T95-99			-0.492 (0.921)	
N	688	145	688	688
No of wars	46	44	46	
Pseudo R^2			0.26	
Log likelihood	-128.21	-38.18	-124.30	

Notes: All regressions include a constant. Standard errors in parentheses. ***, **, * indicate significance at the 1, 5 and 10 percent level, respectively.

We analyzed whether time effects matter by including time dummies in the model. Based on a log likelihood ratio test we cannot reject the hypothesis that the time dummies are zero.[14]

[14] LRT: χ^2 =7.83, 6 restrictions, p=0.251.

Finally, in the last column of Table 8 we use a recently developed correction method for rare events data (King and Zeng, 2001). Our data is characterized by a relatively small number of events (wars), only about seven percent of the observations being characterized by a civil war outbreak. The results presented in King and Zeng (2001) suggest that standard logit estimation underestimates the probability of an event occurring when the events are rare. We used their correction procedure but found the differences between the standard logit results and the rare events corrected results to be negligible. Using the rare events logit procedure, all coefficients on the variables have the same signs and are significant at the same levels. The mean of the predicted probabilities obtained from the rare events logit regression have a mean of 0.072. Thus, we find that the corrected results are very similar to the logit results. We therefore conclude that the use of uncorrected logit regressions is an appropriate method for our empirical analysis.

We examined a number of different model specifications. We found that none of the following geographic and demographic characteristics were significant: forest coverage, population density and the proportion of young men aged 15 to 29.[15]

5. Conclusion and Interpretation

Using a comprehensive data set of large-scale civil conflict over the period 1960-99 we examine the risk of civil war using logit regressions to predict the risk of war in each five-year period. We find that a model which focuses on the constraints facing rebel organizations performs well, whereas objective indicators of grievance add little explanatory power. The model is robust to a range of tests for outliers, redefinitions and alternative specifications.

The level, growth and structure of income are all economically and statistically significant factors in conflict risk. Whether higher income is associated with reduced conflict because it strengthens government finances, because it occupies young males in secondary school, or because of some more general attitudinal effects remains unresolved. That faster growth reduces conflict risk may be because it raises the opportunity cost of joining rebellion, although it could simply generate hope of better times. That primary commodity dependence increases the risk of conflict is consistent with the evident role which primary commodities play as sources of rebel finance. Primary commodity dependence is also associated with poor governance and increased exposure to economic shocks (Auty, 2000), either of which could increase conflict risk. However, we found that a further important risk factor is the size of the diaspora. The case study literature highlights the role of diasporas in financing rebel organizations. Hence, if only by Occam's Razor, this strengthens our interpretation that the route by which primary commodities affect conflict is via rebel finances. At high levels of primary commodity dependence conflict risk declines again: we have interpreted this as due to the strengthening of government finances, and hence deterrence capability.

Despite the attention given to ethnic and religious differences as sources of conflict, we have found no effect of polarization, while greater social fractionalization actually reduces conflict risk. This is consistent with the rebel organization having a disproportionate need for social cohesion in its recruitment, so that rebellion is more difficult in fractionalized societies. We

[15] The proportion of the population living in urban areas was statistically significant when we excluded the geographic dispersion of the population. However, when we included both proxies for the concentration of the population, the geographic dispersion measure remained statistically significant while the proportion of the population living in urban areas was marginally insignificant (p=0.11).

do, however, find that ethnic dominance (defined as the largest group constituting 45-90 percent of the population) moderately increases conflict risk. We have found no effects of either political rights or inequality on conflict risk: hence, we concur that `there is at best a tenuous relationship between war and justice as its motive' (Anderson, 1999, p.11).

We have found that geography matters. Conflict risk does not increase proportionately with population, this being consistent with scale economies in deterrence capability. However, risk is increased by population dispersion and possibly by mountainous terrain.

Overall, these results are consistent with economic models of conflict risk in which the critical parameters are the financial opportunities for rebels, the social and geographic constraints which they face, and the financial capability of the government to provide defense. They are harder to reconcile with accounts of conflict which stress ethnic, religious, political or economic grievances.

References

Alesina, A., S. Oetzler, N. Roubini, and P. Swagel. 1996. 'Political Instability and Economic Growth.' *Journal of Economic Growth* 1:189-211.

Anderson, Mary B. 1999. *Do No Harm*. Boulder, CO: Lynne Reinner Publishers, Inc.

Angoustures, Aline and Valerie Pascal. 1996. 'Diasporas et financement des conflits,' in Francois Jean and Jean-Christohpe Rufin (eds.) *Economie des Guerres Civiles*, Hachette, Paris.

Auty, R.M. 2000. 'How Natural Resources Affect Economic Development.' *Development Policy Review* 18(4):347-64.

Barro, R. J. 1991. 'Economic Growth in a Cross Section of Countries.' *The Quarterly Journal of Economics* 106:407-43.

Barrett, D. B., ed. 1982. *World Christian Encyclopedia.* Oxford: Oxford University Press.

Buchanan, J.M. and R.L. Faith. 1987. 'Secession and the Limits of Taxation: Towards a Theory of Internal Exit.' *American Economic Review*, 77:1023-31.

Collier, P. 1999. 'On the Economic Consequences of Civil War.' *Oxford Economic Papers* 51:168-83.

_____. 2000. 'Rebellion as a Quasi-Criminal Activity.' *Journal of Conflict Resolution* 44: 839-53.

_____.2001. Ethnic diversity: An economic analysis of its implications.' *Economic Policy* 32: 129-66.

Collier, P. and A. Hoeffler. 1998. 'On the Economic Causes of Civil War.' *Oxford Economic Papers*, 50:563-73.

Collier, P., A. Hoeffler, and M. Söderbom. 2001. 'On the Duration of Civil.' Policy Working Paper 2861, World Bank, Washington DC.

Davidson R. and J.G. MacKinnon. 1981. 'Several Tests for model specification in the presence of alternative hypotheses.' *Econometrica* 49:781-793.

Deininger, K., and L. Squire. 1996. A New Data Set Measuring Income Inequality. *World Bank Economic Review* 10:565-91.

Deininger, K. and Squire, L. 1998. 'New Ways of Looking at Old Issues: Inequality and Growth.' *Journal of Development Economics* 57:249-287.

Easterly, W. and R. Levine. 1997. 'Africa's Growth Tragedy: Policies and Ethnic Divisions.' *Quarterly Journal of Economics* 113:1203-49.

_____. 1998. 'Troubles with the Neighbours: Africa's Problem, Africa's Opportunity.' *Journal of African Economies* 7(1):120-42.

Esteban, J.-M., and D. Ray. 1994. 'On the Measurement of Polarization.' *Econometrica* 62(4): 819-51.

_____.1999. 'Conflict and Distribution.' *Journal of Economic Theory* 87:379-415.

Gerrard, A.J.W. 2000. 'What is a Mountain?' Mimeo. DECRG, World Bank.

Gleditsch, Nils Petter, and Havard Hegre. 1997. Peace and Democracy: Three Levels of Analysis. *Journal of Conflict Resolution,* 41:283-310.

Grossman, H. I. 1999. Kleptrocracy and Revolutions. *Oxford Economic Papers* 51:267-283.

Herbst, Jeffrey. 2000. *States and Power in Africa*. Princeton. Princeton University Press.

Hirshleifer, J. 1991. 'The Paradox of Power'. *Economics and Politics* 3:171-200.

Jaggers, Keith and Ted Robert Gurr. 1995. 'Tracking Democracy' s Third Wave with the Polity III Data'. *Journal of Peace Research* 32:469-482.

Kaplan, Robert, D. 2000. *The Coming Anarchy*, New York, Random House.

King, G., and L. Zeng. 2001. ' Logistic Regression in Rare Events Data'*Political Analysis* 9(2):137-163.

Konrad, Kai, A. and Stergios Skaperdas. 1998. 'Extortion'. *Economica*, 65:461-77.

Londregan, J.B. and K. T. Poole. 1996. 'Does High Income Promote Democracy?' *World Politics* 49:1-30.

Long, J. Scott. 1997. *Regression models for categorical and limited dependent variables.* London: Sage Publications.

Mauro, P. 1995. 'Corruption and Growth'. *The Quarterly Journal of Economics* 110:681-712.

Pregibon, D. 1981. Logistic Regression Diagnostics. *The Annals of Statistics* 9:705-24.

Reynal-Querol, M. 2000. '*Religious Conflict and Growth: Theory and Evidence*'. Ph. D. Thesis, London School of Economics and Political Science.

Sachs, J. and A.M. Warner. 2000. 'Natural Resource Abundance and Economic Growth'. In *Leading Issues in Economic Development*, 7th ed. G.M. Meier and J.E. Rauch, Oxford University Press.
Sen, A. 1973. *On economic inequality*. Oxford: Clarendon Press.

The Stockholm International Peace Research Institute. 2000. *Yearbook of World Armaments and Disarmaments.* Oxford: Oxford University Press.

Summers, R., and A. Heston. 1991. The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. The Quarterly Journal of Economics 99:327-68.

USSR, 1964. *Atlas Narodov Mira*, Department of Geodesy and Cartography of the State Geological Committee of the USSR. Moscow.

World Bank, 2000. *World Development Indicators*, Washington DC.

Appendix

A1: A Simple Migration model

Our estimation of migration is based on the following model:

$$dias_{it} = \underset{(0.045)***}{1.163}\, dias_{i,t-1} - \underset{(0.001)**}{0.002}\, GDP_{i,t-1} + \underset{(0.03)}{0.003}\cdot war_{i,t-1} + \underset{(0.002)}{0.003}\cdot T_{80} + \underset{(0.002)***}{0.005}\cdot T_{90} + \underset{(0.008)}{0.013}$$

Where dias denotes diaspora which is measured as the ratio of emigrants in the USA to the total population of the country of origin. The variable "war" is a war dummy, measured at t-1 it takes a value of one if the country experienced a civil war in the previous period. The method of estimation is OLS. The data is measured at the beginning of each decade, i.e. 1960, 1970, 1980 and 1990. The regression includes time dummies, T, which are jointly significant.

Based on this simple migration model we estimated the size of the diaspora at time t.

$$\hat{dias}_{it} = x_{it} \cdot \hat{\beta}$$

For countries which experienced a previous civil war we used these estimated values to correct for a possible endogeneity problem.[16] We replaced a total of 64 observations. For countries which did not experience a civil war we use the actual diaspora data. We took the averages of this corrected diaspora data measured in 1960 and 1970 (1970 and 1980, 1980 and 1990) in order to obtain values for 1965, 1975 and 1985. For 1995 we use the observations measured in 1990.

[16] Here we only consider wars after 1960.

Table A2: Marginal Probabilities

Variable	Coefficient	Mean of X	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
			At the mean	Worst	Best	1% extra growth	10% extra men in school	primary commodity/ GDP = 0.32	primary commodity/ GDP = 0	Ethnic dominance	Maximum fractionalization	Minimum peace	Geographic dispersion = 0
Male secondary schooling	-0.032	44.489	-1.406	-0.032	-4.645	-1.406	-1.722	-1.406	-1.406	-1.406	-1.406	-1.406	-1.406
$(\text{GDP growth})_{t-1}$	-0.115	1.618	-0.186	1.508	-1.660	-0.417	-0.186	-0.186	-0.186	-0.186	-0.186	-0.186	-0.186
Primary commodity exports/GDP	18.937	0.158	2.992	6.060	0	2.992	2.992	6.060	0	2.992	2.992	2.992	2.992
(Primary commodity exports/GDP)^2	-29.443		-0.735	-3.015	0	-0.735	-0.735	-3.015	0	-0.735	-0.735	-0.735	-0.735
Social fractionalization	-0.0002	1790	-0.376	-0.004	-1.465	-0.376	-0.376	-0.376	-0.376	-0.376	-1.465	-0.376	-0.376
Ethnic Dominance (45-90%)	0.670	0.439	0.294	0.670	0	0.294	0.294	0.294	0.294	0.670	0.294	0.294	0.294
Peace Duration	-0.004	347.5	-1.286	-0.004	-2.19	-1.286	-1.286	-1.286	-1.286	-1.286	-1.286	-0.004	-1.286
Ln Population	0.768	30,500,000	13.230	16.049	9.136	13.230	13.230	13.230	13.230	13.230	13.230	13.230	13.230
Geographic dispersion	-2.487	0.602	-1.497	0.000	-2.415	-1.497	-1.497	-1.497	-1.497	-1.497	-1.497	-1.497	0
Constant	-13.073		-13.073	-13.073	-13.073	-13.073	-13.073	-13.073	-13.073	-13.073	-13.073	-13.073	-13.073
$X \cdot \hat{\beta}$			-2.043	8.160	-16.312	-2.273	-2.359	-1.255	-4.300	-1.667	-3.132	-0.761	-0.546
W			0.130	3497.6	0.000	0.103	0.095	0.285	0.014	0.189	0.044	0.467	0.579
Probability			0.115	1.000	0.000	0.093	0.086	0.222	0.013	0.159	0.042	0.318	0.367

Table A3: Marginal Probabilities–Alternative Specification

Variable	Coeffi-cient	Mean of X	Mean	Double mean income	Plus one st. dev.	Plus two st. dev.	Minimum income	Maximum income
ln GDP per capita	-0.950	3957	-7.869	-8.528	-8.535	-8.923	-5.133	-9.359
$(\text{GDP growth})_{-1}$	-0.098	1.530	-0.150	-0.150	-0.150	-0.150	-0.150	-0.150
Primary commodity exports/GDP	16.773	0.155	2.600	2.600	2.600	2.600	2.600	2.600
$(\text{Primary commodity exports/GDP})^2$	-23.800	0.024	-0.572	-0.572	-0.572	-0.572	-0.572	-0.572
Social fractionalization	0.000	1812.780	-0.453	-0.453	-0.453	-0.453	-0.453	-0.453
Ethnic dominance (45-90%)	0.480	0.437	0.210	0.210	0.210	0.210	0.210	0.210
Peace duration	-0.004	348.210	-1.323	-1.323	-1.323	-1.323	-1.323	-1.323
ln Population	0.511	30600000	8.799	8.799	8.799	8.799	8.799	8.799
Geographic Dispersion	-0.992	0.600	-0.595	-0.595	-0.595	-0.595	-0.595	-0.595
Constant	-3.438		-3.438	-3.438	-3.438	-3.438	-3.438	-3.438
$X \cdot \hat{\beta}$			-2.791	-3.450	-3.457	-3.845	-0.055	-4.281
W			0.061	0.032	0.032	0.021	0.947	0.014
Probability			0.058	0.031	0.031	0.021	0.486	0.014

A3: Data Sources

Diaspora

In order to proxy the size of the diaspora we used US immigration data from the US Bureau of the Census. To capture the relative size of the population we divided the US immigration figures by the total population of the country of origin. Data source: http://www.census.gov/population

Ethnic dominance

Using the ethno-linguistic data from the original data source (Atlas Narodov Mira, 1964) we calculated an indicator of ethnic dominance. This variable takes the value of one if one single ethno-linguistic group makes up 45 to 90 percent of the total population and zero otherwise. We would like to thank Tomila Lankina for the translation of the original data source.

Forest Coverage

We used the FAO measure of the proportion of a country's terrain which is covered in woods and forest. Source: http://www.fao.org/forestry

GDP per capita

We measure income as real PPP adjusted GDP per capita. The primary data set is the Penn World Tables 5.6 (Summers and Heston 1991). Since the data is only available from 1960-92 we used the growth rates of real PPP adjusted GDP per capita data from the World Bank's World Development Indicators 1998 in order to obtain income data for the 1990s.

Geographic Dispersion of the Population

We constructed a dispersion index of the population on a country by country basis. Based on population data for 400km^2 cells we generated a Gini coefficient of population dispersion for each country. A value of 0 indicates that the population is evenly distributed across the country and a value of 1 indicates that the total population is concentrated in one area. Data is available for 1990 and 1995. For years prior to 1990 we used the 1990 data. We would like to thank Uwe Deichman of the World Bank's Geographic Information System Unit for generating this data. He used the following data sources: Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); and World Resources Institute (WRI). 2000. Gridded Population of the World (GPW), Version 2. Palisades, NY: IESIN, Columbia University. Available at http://sedac.ciesin.org/plue/gpw.

Mountainous terrain

The proportion of a country's terrain which is mountainous was measured by John Gerrard, a physical geographer specialized in mountainous terrain. His measure is not only based on altitude but takes into account plateaus and rugged uplands. The data is presented in Gerrard (2000).

Population
Population measures the total population, the data source is the World Bank's World Development Indicators 1998.

Primary commodity exports/GDP

Following Sachs and Warner (2000) we measure the abundance of natural resources by the ratio of primary commodity exports to GDP. The data on primary commodity exports as well as GDP was obtained from the World Bank. Export and GDP data are measured in current US dollars.

Social, ethnolinguistic and religious fractionalization

We proxy social fractionalization in a combined measure of ethnic and religious fractionalization. Ethnic fractionalization is measured by the ethno-linguistic fractionalization index. It measures the probability that two randomly drawn individuals from a given country do not speak the same language. Data is only available for 1960. In the economics literature this measure was first used by Mauro (1995). Using data from Barrett (1982) on religious affiliations we constructed an analogous religious fractionalization index. Following Barro (1997) we aggregated the various religious affiliations into nine categories: Catholic, Protestant, Muslim, Jew, Hindu, Buddhist, Eastern Religions (other than Buddhist), Indigenous Religions and no religious affiliation.

The fractionalization indices range from zero to 100. A value of zero indicates that the society is completely homogenous whereas a value of 100 would characterize a completely heterogeneous society.

We calculated our social fractionalization index as the product of the ethno-linguistic fractionalization and the religious fractionalization index plus the ethno-linguistic or the religious fractionalization index, whichever is the greater. By adding either index we avoid classifying a country as homogenous (a value of zero) if the country is ethnically homogenous but religiously divers, or vice versa

Peace Duration

This variable measures the length of the peace period since the end of the previous civil war. For countries which never experienced a civil war we measure the peace period since the end of World War II.

War Starts

We use mainly the data collected by Singer and Small (1984, 1994). We would like to thank Nicholas Sambanis for extending and updating this data set for us.