

Conflict determinants in Africa

J. Paul Dunne and Nan Tian

J. Paul Dunne is Professor of Economics, School of Economics, University of Cape Town, Cape Town, South Africa. The corresponding author, he may best be reached at john.dunne@uct.ac.za. Nan Tian is Researcher, Arms Transfers and Military Expenditure Programme, Stockholm International Peace Research Institute, Stockholm, Sweden. He may best be reached at nan.tian@sipri.org.

Abstract

This article considers the determinants of conflict in Africa. It revisits the greed versus grievance debate to consider the specific regional context and changing nature of conflict in Africa. This is a literature that has grown rapidly in economics and political science, but some recent developments in modeling and conceptualization are providing important new contributions. The article uses the zero-inflated ordered probit technique that deals with the problem of excess zeros in datasets, revisits the definition of conflict, and improves upon some proxy measures. It also considers the substantive as well as statistical significance of the variables. Changes in the technique used provide more support for the influence of grievance terms than given credit for with the usual probit model approach. Both greed and grievance determine conflict in Africa.

Civil war has been commonplace for the past 60 years, but until fairly recently it received little attention from economists. Civil war is not just common; it is also persistent and lasting longer, decade after decade (Fearon, Kasara, and Laitin, 2007). Blattman and Miguel (2010) estimated that, since 1960, some 20 percent of all countries have experienced at least 10 years of civil conflict, often devastating them culturally, politically, and economically. Collier, *et al.* (2003) suggested that the destructive forces could be large enough to explain the income gap between the poorest and richest nations. One could almost see civil war as reversing development, diverting resources from productive activities to destruction and having both, devastating direct costs and opportunity costs from the loss of productive resources (Collier, *et al.*, 2003). The actual and potential costs make it important to understand why conflicts start, and the contribution by Collier and Hoeffler (2004), which sought to test two competing theoretical hypotheses concerning the determinants of intrastate armed conflict—opportunity, or “greed”, versus grievance—has led to a large empirical literature. Their finding of overwhelming support in favor of the view that rebellion is motivated by opportunity is generally accepted but has become rather more nuanced (Blattman and Miguel, 2010).

As researchers started to accept the general framework, they also examined other potential determinants that had not already been considered. Nowhere is this move beyond greed or grievance more evident than in quantitative studies of conflict prevalence in Africa where the imposition of artificial state borders, living in “bad neighborhoods”, and warmer temperatures (increasingly so, in the face of climate change)

have come to take center stage as explanatory variables of interest in the econometric models employed in these studies (Hendrix and Glaser, 2007; Burke, *et al.* 2009).

A number of developments have led to a point where there is some value to be gained from revisiting the debate. First, there are obviously more years of data available, more economic shocks, and more conflicts. Second, there have been significant improvements in the operationalization of difficult-to-measure indicators of greed and of grievance (e.g., income inequality, ethnic divisions). And third, there have been developments in the estimation methods available for analysis, in particular the recognition that simple probit, or logit, models do not perform well in situations with a large number of zeroes in the dependent variable, a likely case for civil conflict as, fortunately, many country-year observations are zero (i.e., peace; see Dunne and Tian, 2017).

A brief review of the determinants of civil war literature and the greed–grievance debate is provided in the next section, followed by a discussion and outline of the estimation procedure used, in particular the zero-inflated ordered probit (ZiOP) model. The section thereafter presents variable construction, the data used, and some descriptive statistics, followed by empirical estimates of a greed–grievance model using the usual methods as against the ZiOP model, and with robustness checks. The final section offers conclusions and discussion of policy implications.

Causes of civil conflict

A range of theoretical perspectives inform the analysis of civil wars. These reflect the interdisciplinary nature of the research

and the relatively late involvement of economists. Political scientists focused upon grievance-related determinants of conflict, with theories emphasizing how modernization could lead to disruption of social order, with social and economic change causing the breakdown of social cohesion and alteration of perceptions. A formalization of this perspective was provided by political rational choice theories. These focused on the role of political repression, failing institutions, political transitions, and informational problems, which together with a failure to redress grievances—economic or political—can lead to conflict.

An alternative was provided by constructivist theories, which focused on the social construction of identity, rather than accepting identity as some fixed attribute. Here, political mobilization leads to civil violence, with leaders constructing ethnic and social identity in ways that benefit themselves (Sambanis, 2002).

In contrast, the focus of economists was on “greed”, or opportunity-based, determinants of conflict. Grossman (1991) modeled rebellion as an industry, while Hirschleifer (1995) suggested it was possible for rational agents to misperceive opportunities and grievances because of asymmetric information. This perspective suggests that civil conflict onset is linked to the possibility (and ability) of insurgents to make a profit—the greed hypothesis—rather than the result of grievances (Dunne and Coulomb, 2008; Skaperdas, 2008).

Collier and Hoeffler (2004) provided an empirical analysis of these competing hypotheses, suggesting that while political grievances are universal, economic incentives are not, and so are often decisive in the start of conflict. The probability of rebel victory depends on the ability of the incumbent to defend, which is determined by technology (the technology may also be available to the rebels, but is limited) and by military expenditure, to which the rebels do not have access.¹ Factors that influence opportunity (such as finance, cost of rebellion, and military advantage) were statistically significant in determining civil war, while most proxies of grievance (ethnic fractionalization, inequality, and democracy) were insignificant, although population size had an effect and time seemed to heal the damaging effects of conflict. This finding—that opportunity explained conflict risk—supported the economic interpretation of rebellion as motivated by greed (Collier and Hoeffler 2004; 2007).

Around the same time, Fearon and Laitin (2003) developed a different model, a game of insurgency where the size of a rebellion is influenced by government effort and the scale of the initial rebellion. They, too, found that political grievance had little explanatory power, but that state institutional capacity was significant, suggesting that wars are caused by countries having weak institutions. Yet they differed from Collier and Hoeffler

This article revisits the greed versus grievance debate to consider the specific regional context and the changing nature of conflict in Africa as recent developments in statistical modeling and conceptualization are providing important new contributions. In particular, the article use a modeling technique (zero-inflated ordered probit) that deals with the statistical problem of excess zeros in the dataset, revisits the definition of conflict, and improves upon some proxy measures. It also considers the substantive as well as statistical significance of the variables. The results provide more support for the influence of grievance terms than ordinarily found with the usual ordered probit model.

(2004) in the interpretation of GDP per capita (reflecting state capacity rather than as an opportunity cost), how civil wars were coded, and using annual data rather than five-year data averages. The two papers (F&L and C&H, in the tables below) had a major bearing on research and debate and led to a large literature that has advanced our understanding on civil conflict telling us what we do know, as well as what we do not (Blattman and Miguel, 2010).

While the general consensus in the literature—that the motivations of greed outweigh those of grievance in explaining civil war onset—remained, the literature continued to develop and improve in a number of areas. First, political scientists questioned the apparent lack of significance of variables that are proxies for objective grievance.² This led to efforts being made to improve measurement and to obtain better proxies. This included improvement of natural resource data, better measurement of grievances, such as measures of inequality and the consideration of horizontal and vertical inequality, and better measures of weak institutions (Lujala, Gleditsch, and Gilmore, 2005; Wucherpfennig, *et al.*, 2011). Second, some attempts have been made at improving causal identification. The potential endogeneity of GDP to conflict led to the use of rainfall as an instrument, given that it may affect agrarian economies’ output, but not conflict. Other attempts have used price shocks and trade shocks in a similar manner. The identification problem remains an issue, mostly due to difficulties in finding appropriate instruments (Blattman and Miguel, 2010; Miguel, Satyanath, and Sergenti, 2004). Third, some attempts have been made to consider possible spillover effects of conflicts, creating conflicts in other countries, with the feedback of refugees keeping conflicts going (Salehyan and Gleditsch, 2006; Dunne and Tian, 2014). Fourth, questions have been raised about measures of conflict and violence. In the past, war tended to be defined as an event in which there were more than 1,000 battle-related deaths (and peace defined as less than this). Initially, this definition was developed for interstate conflicts and then continued in use, even after the

focus shifted to civil conflicts in the post-cold war world. Eventually, this was deemed unsuitable, and an added definition of conflict (more than 25 battle-related deaths), was created.³

This article engages with a fifth concern, the estimation method used. An ordered probit model with a zero–one dependent variable for conflict has generally been used, but this includes a lot of zeros (peace years) in the dataset, and these zeroes are unlikely to all stem from the same data generation process. An observation of a year of peace for a country that is in and out of war surely is different to one for a country that is generally at peace. A zero value in a particular year for Botswana, for example, is rather different to one for the DR Congo (Bagozzi, *et al.*, 2015).

Data and units of measurement

To operationalize a greed–grievance empirical model, data for a range of variables were collected, following developments in the literature. Two sets of income variables (real GDP in purchasing power parity terms and its per capita growth rate) were taken from the World Bank and Penn World Tables 8.0, as well as the degree of urbanization (the proportion of a country’s population living in an urban environment), and life expectancy (in years).⁴ The percentage of mountainous terrain in a given country was also considered, as an indicator of military accessibility or safe havens for rebels.⁵

Natural resource dependence was proxied by the percent share of primary commodity exports in GDP⁶ but, given the ongoing debates on the measure of natural resource dependence and the type of commodities used, three additional measures were considered. First and second, annual oil production in metric tons and oil exports greater than one-third of total exports were collected as proxies for oil abundance and dependence, respectively.⁷ And, third, to distinguish fuel and nonfuel minerals from other primary commodities, a mineral dependence variable was created. A country was considered mineral dependent if its mineral exports constituted 25 percent or more of total tangible exports. Collectively, these variables are our opportunity variables (see Table 1).

Our grievance variables are, for the most part, common to those identified by Fearon and Latin (2003) and Collier and Hoeffler (2004). They fall into three groups: (1) ethnic and/or religious hatred, (2) political repression, or freedom, and (3) horizontal income inequality. As to the first, the most commonly chosen indicator to test for any link between ethnicity and civil conflict is ethnic fractionalization.

Table 1: Descriptive statistics, means

	<i>Full sample</i>	<i>Always 0</i>	<i>Not always 0</i>	<i>Civil war</i>	<i>No civil war</i>
<i>(See text for units of measurement.)</i>					
<u>OPPORTUNITY VARIABLES</u>					
GDP/capita	7,931	14,069	3,311	3,172	8,699
GDP/capita growth	1.8	2.2	1.6	1.0	1.9
Urbanization rate	46.9	56.0	39.7	40.6	47.9
Life expectancy	61.6	66.2	58.0	59.4	62.0
Mountains	16.38	14.93	18.11	23.16	15.33
Primary commodity exports/GDP	15.6	17.8	13.9	10.9	16.4
Oil production	17,000	13,7000	19,300	19,100	16,700
Oil exports	18.7	15.5	20.8	16.8	18.9
Mineral dependence	49.3	41.5	54.5	55.5	48.4
<u>GRIEVANCE VARIABLES</u>					
Ethnic fract.	63.0	52.1	69.9	77.5	60.1
Ethnic dominance	47.0	48.3	46.7	54.9	45.7
Religious frac.	36.5	36.1	36.6	0.36	0.37
Polity IV	1.13	3.84	−0.73	0.97	1.30
LDG (see Notes)	0.056	0.024	0.081	0.142	0.042
NHI (see Notes)	1.189	1.064	1.278	1.398	1.155
PHI (see Notes)	1.201	1.086	1.287	1.224	1.197

Notes: LDG = largest discriminated (against) ethnic group; NHI = negative horizontal inequality (relative gap between mean national income and income level of the poorest group); PHI = positive horizontal inequality (relative gap between mean national income and income level of the richest group).

Measurement of this, in Table 1, is taken from Collier and Hoeffler (2004), with ethnic dominance measured as a binary variable, taking on the value of 1 if the largest ethnic group in a country amounts to 45–90 percent of the population, and religious fractionalization similarly measured.⁸ Regarding the second group, data from the Polity IV database was used to measure the degree of political rights, with the variable ranging from −10 (high autocracy) to +10 (high democracy). In the regressions to follow, we include a squared term to allow for nonlinear effects (Hegre, *et al.*, 2001). And third, Buhaug, Cederman, and Gleditsch (2014) found that certain indices of horizontal income inequality and political discrimination (LDG, NHI, and PHI in Table 1) performed much better than conventional indicators and these are used in robustness checks for ethnopolitical and economic grievance.⁹

The control variables included in our regressions are the standard ones found in the literature (e.g., population and the cold war period; not shown in Table 1 but shown in Tables 2 and 3). The dependent variable, conflict prevalence, takes on three values, namely 0 for all peace year observations, 1 for “minor” conflict years with combat deaths ranging between 25–999 people, and 2 for “major” civil wars with annual battle deaths with 1,000 or more people.

Table 1 shows that the always zero or “complete peace” group has higher GDP per capita (level and growth), greater urbanization, better life expectancy, and more political freedom than the “not always zero” group and also exhibits lower levels of ethnic and religious fractionalization and income inequality. Correlations suggested some association between the income and inequality variables and the likelihood of a country being completely peaceful versus having some experience of conflict. In episodes of civil conflict, GDP per capita (level and growth), urbanization, life expectancy, and political freedom all are lower than in times of peace, while ethnic divisions, income inequality, and substantial amounts of rough terrain are higher for civil war episodes. Interestingly, primary commodity exports (share of GDP) is on average lower for civil war years.

Greed versus grievance revisited

Estimating the probability of civil conflict using an ordered probit gave the results in Table 2. Column (1) gives the results with the ethnolinguistic fractionalization variable used by Fearon and Laitin (2003) and column (2) when this is replaced by the Collier and Hoeffler (2004) measure. Column (3) gives the results when, following Buhaug, Cederman, and Gleditsch (2014), other ethnic discrimination and income inequality measures are introduced instead. (These are denoted as F&L, C&H, and BC&G, respectively.) Taking the opportunity variables first, all six signs for GDP and per capita GDP growth are negative, suggesting that higher income moderates the likelihood of civil war. Primary commodity exports as a share of GDP exert a nonlinear effect on conflict prevalence, first decreasing and then increasing,¹⁰ and mountainous terrain increases civil war risk. The magnitudes, signs, and statistical significance for all these variables are very similar across the three specifications.

As for the grievance variables, political freedom as captured by the Polity IV index is statistically insignificant, and the results for ethnicity vary, conditional on the choice of the ethnic fractionalization variable. When the Fearon and Laitin (2003) variable is used it is significant, but insignificant for the Collier and Hoeffler (2004) variable. Ethnic dominance is significant and positive, but religious fractionalization, while positive, is only significant for model (1). Column (3) shows LDG, the

largest discriminated against ethnic group, as a proxy for ethnic and political inequality, to be positive and highly significant. NHI, the measure of negative horizontal inequality, is positive and statistically significant at the 1 percent level as well (suggesting that African countries with one or more ethnic groups that are radically poorer than the national average have a higher risk of conflict onset), and PHI, the positive horizontal inequality, to be negative and significant.¹¹ The log of population control variable is positive and significant for all specifications, but the cold war dummy variable is negative and significant only for specification (3).

In most analyses of the determinants of civil conflict, an ordered dependent variable is used, in which a given country-year is assigned a value of 0 for peace and a value of 1 when violence between the state and another side reaches a given threshold, thereby classifying it as a civil war. Since peaceful years dominate conflict years, a very large number of zero observations are in the dataset. These can reflect rather different states of peace, however, namely, one where structural and societal forces ensure a zero probability of civil conflict regardless of greed or grievance incentives or another that captures a mere interlude in fighting and a high probability of returning to conflict.

In the first group will often be states such as Botswana, which can be labeled as “complete-peace.” The second group contains states in regions such as Central, West, or East Africa and can be labeled as “incomplete-peace.” (Boulding, 1978, might call the groups “stable peace” and “unstable peace”). The main difference between the first and second case of zeroes is that while the probability of transition into war for the first type is zero, the probability for the second group is not zero. In the latter case, incentives resulting from opportunity and/or grievance can induce violent conflict.

Given the high proportion of heterogeneous zeroes in the analysis, using conventional probit, or logit, models may not be appropriate tools for statistical inference and can potentially give biased estimates (Bagozzi, *et al.*, 2015). In such cases, a more satisfactory estimation method is the split-population or two-part model proposed by Harris and Zhao (2007) and Vance and Ritter (2014). This is typically done in the form of zero-inflated models or, in our case, a zero-inflated ordered probit (ZiOP) model, where estimations follow two stages. The first is a selection or *inflation equation*, which splits the observations into two processes, each potentially having different sets of explanatory variables. In the context of civil war prevalence, zero observations in process 0 ($w_i=0$) include inflated zeroes, consistent with countries that never experience civil conflict (e.g., Botswana), while zero observations in process 1 ($w_i=1$) includes cases for which the probability of

transitioning into a civil conflict is not zero, even if civil war casualties have not reached the lower bound (or limit) of 1,000 battle-related deaths. The binary variable, w_i , thus indicates the split between process 0 (with $w_i = 0$ for no war) and process 1 (with $w_i = 1$ for war). A second stage estimates the ordered probit *outcome equation*, conditioned on the first stage. A fuller exposition of the model is provided in the Appendix.¹²

Compared to standard probit or logit models, the ZiOP model allows more accurate estimates to be obtained but it should be noted that the usefulness of the model (i.e., unbiased estimates) declines when the size of the split in the sample population becomes very big or very small, leading to biased results.¹³ Bagozzi, *et al.* (2015) suggest that this becomes an issue when there are less than 10 percent or greater than 90 percent of zero observations. In our case, the zero observations comprise about 76 percent of the dataset.

For the Fearon and Laitin (2003) measure of ethnic division, the results of two specifications of the ZiOP model are given in Table 3. In the first, the inflation equation is limited to GDP (level and growth), political freedom (Polity IV), and ethnic fractionalization as these factors promote interest compatibility between the state and its citizens, which in turn influences the probability that a country is in the always zero group and always experiences peace. That said, to ensure that the ZiOP estimates are not driven by choice of variables, a second specification includes all covariates in the outcome equation in the inflation equation as well. This second specification is used to check that the results do not change markedly when the specification of the inflation/selection equation changes. This is to show that the researchers have not simply searched for a

Table 2: Ordered probit of civil war prevalence, 1960–2013

	(1) [F&L]	(2) [C&H]	(3) [BC&G]
<u>OPPORTUNITY VARIABLES</u>			
log, real GDP	-0.024 (0.052)	-0.137** (0.052)	-0.202** (0.050)
Real GDP per capita growth	-2.496** (0.523)	-2.455** (0.520)	-2.492** (0.527)
Prim. exp./GDP	-5.329** (0.966)	-5.601** (0.978)	-4.091** (0.963)
Prim. exp./GDP squared	7.801** (1.585)	9.045** (1.595)	5.908** (1.597)
log, mountains	0.054* (0.028)	0.118** (0.030)	0.062* (0.028)
<u>GRIEVANCE VARIABLES</u>			
Polity IV	-0.015 (0.032)	-0.018 (0.032)	0.001 (0.031)
Polity IV squared	0.001 (0.005)	0.003 (0.005)	-0.005 (0.004)
Ethno frac. (F&L)	6.022** (0.998)		
Ethno frac. squared (F&L)	-5.529** (0.934)		
Ethno frac. (C&H)		0.011 (0.008)	
Ethno frac. squared (C&H)		-0.001 (0.001)	
Ethnic dominance	0.210* (0.086)	0.292* (0.119)	
Religious frac.	0.967** (0.301)	0.218 (0.275)	
LDG (see Notes)			1.264** (0.168)
NHI (see Notes)			0.859** (0.125)
PHI (see Notes)			-0.172* (0.078)
<u>OTHER VARIABLES</u>			
log, population	0.340** (0.066)	0.413** (0.068)	0.514** (0.679)
Cold war period	-0.024 (0.097)	-0.008 (0.096)	-0.043** (0.104)
Observations	1,519	1,519	1,542
Log likelihood	-941.8	-944.3	-901.8
AIC	1,913.65	1,918.67	1,835.67

Notes: Dependent variable = conflict prevalence (0, 1, or 2); AIC = Akaike Information Criterion; Standard errors in parentheses; significance levels: ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$; LDG = largest discriminated (against) ethnic group; NHI = negative horizontal inequality (relative gap between mean national income and income level of the poorest group); PHI = positive horizontal inequality (relative gap between mean national income and income level of the richest group).

specification that “works”.¹⁴

Looking in Table 3, then, at the first stage or *inflation equation* for specification (1), the results show the GDP variables (level and growth) with a statistically significant negative effect on the likelihood of a country-year not being among the always-zero or peace group and then experiencing any level of civil violence. Additionally, political freedom, measured by the Polity IV index, has the usual nonlinear effect of first increasing the likelihood of civil conflict and then decreasing it past a certain point. Ethnicity also plays an

important role in the different observed zeroes. A higher degree of fractionalization makes it more likely that a country will experience conflict.

The *outcome equation* of the ZiOP specification (1) in Table 3 can then be directly compared to the conventional ordered probit specification (1) in Table 2. While the signs are generally consistent, there are substantial differences as well. Among the opportunity variables, for instance, the coefficient for real GDP under ZiOP is over 10 times larger and the presence of mountainous terrain no longer is statistically significant.

As for the grievance terms, political freedom now is a significant predictor of civil war prevalence and its coefficient estimate is considerably larger.

Moreover, the effect is not of the usual “inverse U” shape but decreases throughout. This is an interesting finding. It suggests that *any* improvement in political freedom lowers the likelihood of civil war (albeit with diminishing effect). Fractionalization (ethnic and religious) remains significant, as before.

Compared to the standard ordered probit model, the ZiOP estimates also have lower standard errors and a lower Akaike Information Criterion (AIC), suggesting that the model better fits the data. As suggested by Cameron and Trevidi (2010), all regressions were estimated using robust standard errors. Again, note that the proportion of zero observations in the sample, at 76.3 percent, falls within the accepted band of 10 to 90 percent (Bagozzi, *et al.*, 2015).

To consider the robustness of our results, a number of alternative ZiOP specifications were estimated. Adding

Table 3: Probit versus ZiOP regressions of civil war prevalence, 1960–2013

	(1) [F&L]		(2) [F&L]	
	<i>Outcome</i>	<i>Inflation</i>	<i>Outcome</i>	<i>Inflation</i>
<u>OPPORTUNITY VARIABLES</u>				
log, real GDP	-0.249** (0.078)	-0.375** (0.095)	-0.269** (0.072)	-1.247** (0.250)
Real GDP/cap. growth	-1.779** (0.669)	-1.722* (0.876)	-3.148** (0.582)	-1.086 (1.535)
Prim. exp./GDP	-8.574** (1.518)		-6.652** (1.366)	-9.256** (2.785)
Prim. exp./GDP squared	12.536** (2.709)		9.957** (2.450)	5.872† (3.617)
log, mountains	0.033 (0.042)		0.341** (0.040)	0.482** (0.067)
<u>GRIEVANCE VARIABLES</u>				
Polity IV	-0.053** (0.014)	0.060* (0.024)	-0.070† (0.040)	-0.013 (0.033)
Polity IV squared	-0.015** (0.003)	-0.008† (0.004)	-0.011* (0.005)	0.036** (0.006)
Ethno fract.	6.882** (1.378)	0.840* (0.380)	2.609* (1.234)	-2.411 (4.334)
Ethno fract. squared	-6.646** (1.265)		-0.284 (1.215)	21.884** (3.509)
Ethnic dominance	0.735** (0.131)		0.686** (0.128)	-3.481** (0.496)
Religious frac.	1.076* (0.510)		-0.196 (0.378)	6.634** (1.607)
<u>OTHER VARIABLES</u>				
log, population	-0.345** (0.117)	1.220** (0.168)	4.126** (0.432)	4.125** (0.432)
Cold war period	0.439** (0.145)		0.445** (0.121)	3.753** (0.491)
Constant	-	-11.212** (1.469)	-	22.239** (4.046)
Observations		1,519		1,519
Log likelihood		-875.63		-814.56
AIC		1,795.26		1,631.85

Notes: AIC = Akaike Information Criterion; Dependent variable: Conflict prevalence; Standard errors in parentheses; Significance levels: ** p < 0.01, * p < 0.05, † p < 0.1.

horizontal income inequality and ethnic discrimination in place of ethnic dominance and religious fractionalization and replacing Fearon and Laitin’s (2003) ethnic fractionalization measure with Collier and Hoeffler’s (2004) gave results consistent with Table 3, with the ZiOP model preferred to the ordered probit model in almost all instances.¹⁵ Other tests included replacing primary commodity exports with either mineral dependence, oil production, or oil exports, replacing the Polity IV index with the Freedom House measure, democracy, and autocracy dummies, and substituting income variables with the urbanization rate and life expectancy. The results were fairly robust, with primary commodity dependence increasing civil war risk, and democracy, political freedom, and higher urbanization decreasing civil war risk.

Ward, Greenhill, and Bakke (2010) remind us that

coefficients' statistical significance does not necessarily mean that models predict well, an important concern given the influence of some of the literature's results on policy formulation. To evaluate the predictive power of our models, the receiver operating characteristics (ROC) statistic was used, which takes the estimated probabilities and compares them to the actual values of the conflict variable. Using different thresholds, this finds the number of correctly classified/predicted observations.¹⁶ The ROC can range from 0.5 (a nonpredictive model, no better than chance) to 1.0 (perfect prediction). Since we used ordered probit models, the ROC scores needed to be computed for values of 1 and 2 ("minor" and "major" conflict). As shown in the last two rows of Table 4, our ZiOP model (for the Fearon and Laitin run) resulted in larger ROC scores, namely 0.766 as against 0.716 for the standard ordered probit, when the outcome variable was 1, and 0.872 as against 0.808 when outcome variable was 2.¹⁷ This indicates that with the same specifications, our ZiOP model predicted civil conflict better than the ordered probit.

Another concern raised by Ward, Greenhill, and Bakke (2010) is that variables may be statistically significant and yet not contribute much to a model's predictive power. This can be evaluated by deleting one independent variable at a time and measuring the effect the deletion had on predictive power (that is, the *change* in ROC). Table 4 presents these results. For example, when excluded from the ZiOP model, the Polity IV only decreases its predictive power from 0.766 to 0.763 (a decrease of 0.003) if outcome variable equals 1, and from 0.872 to 0.867 (a decrease of 0.005) if the outcome variable equals 2. Although statistically significant, the Polity IV variable does not appear to provide a substantive contribution to the model.¹⁸

Conclusion

This article revisits the greed–grievance debate within the context of fragility, using a data set of 33 African countries for the period 1960 to 2013. This seemed justified for a number of reasons: the existence of more years of data including more economic shocks and more conflicts, the significant

Table 4: Predictive power and statistical significance, probit versus ZiOP

	Ordered Probit		Zero-inflated Ordered Probit		p-values	
	Δ ROC if Outcome = 1	Δ ROC if Outcome = 2	Δ ROC if Outcome = 1	Δ ROC if Outcome = 2		
<u>OPPORTUNITY VARIABLES</u>						
log, real GDP	0.645	-0.002	-0.001	0.001	-0.007	-0.001
RGDPPPC growth	0.000	-0.003	-0.028	0.008	-0.001	-0.004
Pri exports/GDP	0.000	-0.031	0.003	0.000	-0.020	0.003
Pri exports/GDP squared	0.000	-0.021	-0.001	0.000	-0.014	0.002
log, mountains	0.051	-0.005	-0.007	0.420	-0.013	-0.019
<u>GRIEVANCE VARIABLES</u>						
Polity IV index	0.636	0.004	-0.001	0.000	-0.003	-0.005
Polity IV index squared	0.748	0.002	-0.001	0.000	-0.002	-0.008
Ethno fraction. (F&L)	0.000	-0.033	-0.022	0.000	-0.011	-0.004
Ethno fraction. squared (F&L)	0.000	-0.031	-0.021	0.000	-0.013	-0.003
Ethnic dominance	0.015	0.014	-0.011	0.000	-0.004	-0.020
Religious fraction.	0.001	-0.003	-0.09	0.035	-0.009	-0.005
<u>OTHER VARIABLES</u>						
log, population	0.000	0.006	-0.010	0.003	-0.004	0.000
Cold war period	0.803	-0.002	0.002	0.420	0.012	-0.010
Sum		-0.105	-0.107		-0.089	-0.074
ROC AUC if outcome = 1			0.716			0.766
ROC AUC if outcome = 2			0.808			0.872

Notes: ROC = Receiver operating characteristics; AUC = Area under the curve.

improvements in the operationalization of difficult-to-measure indicators of grievance (i.e., income inequality, ethnic divisions), and the development of a new estimation method that seems well suited to the subject. Estimations using the standard ordered probit technique do not account for the heterogeneous zeroes in the dataset, and an alternative, zero-inflated, model is used that separates out observations of countries with almost no probability of conflict from those of other countries.

The two main results are the following. First, unlike much of the earlier literature, civil war risk is not wholly dominated by greed (or opportunity); the grievance terms are statistically significant. It appears that the matter is not one of a disjunctive "greed *or* grievance," but one of a conjunctive "greed *and* grievance." Second, our zero-inflated ordered probit (ZiOP) models perform better statistically than do the standard probit models and better account for observable and latent factors that

produce different types of peace observations. These results suggest that the standard ordered probit technique results in biased estimates, giving greater weight to opportunity over grievance variables. This has led to most empirical work finding opportunity variables as the main determinant of civil conflict (the “disjunctive” result).

As one takes a deeper look at what type of country is mostly associated with the always zero or “complete peace” group, the answer often is higher-income countries. By not distinguishing the different types of zeroes, the standard ordered probit gives a likelihood of war calculation that includes countries conditioned to not experience war. These countries’ main attribute is higher income, and income variables thus are estimated with greater emphasis and significance, crowding out the grievance variables’ explanatory power. In contrast, using a zero-inflated probit model and splitting the estimation process into two stages, opportunity and grievance variables are given equal emphasis, which makes it clear that both greed and grievance matter, and both with substantial explanatory power in predicting civil war risk.

Clearly, economic factors are important in determining conflict prevalence, but so are grievances, and this is clearer when the lower probability of higher income/peaceful countries is considered. In postwar situations, it is important to study the causes of the conflicts with some care, both in terms of greed and grievance factors, and to deal with the underlying problems, rather than believing that general prescriptive policies will suffice (Brauer and Dunne, 2012).

Notes

We are grateful to the African Development Bank for support and to an anonymous referee for helpful suggestions. All remaining errors are ours.

1. Dunne and Tian (2017) provide more detail on these studies.
2. In a recent contribution, Buhaug, Cederman, and Gleditsch (2014) argued that the lack of significance had to do with the poor proxy variables used in previous research. They showed that better proxies indicate that grievances do matter.
3. A further development saw Besley and Persson (2010, 2014) create a nonbinary ordinal measure of civil violence, with 0 as the value for peace, 1 for civil repression, and 2 for large-scale civil conflict with more than 1,000 battle deaths. New datasets are allowing more consistent and detailed information to be used, such as the data set of global instances of political violence (<http://ucdp.uu.se/ged/>).
4. Sourced from the World Bank, the degree of urbanization can also be thought of as a measurement of geographic dispersion: The greater the urbanization, the lower the geographic dispersion. All income figures are adjusted for purchasing power parity (PPP). Male secondary school enrollment was not

used in the estimations due to poor and incomplete data.

5. Pickering (2011) criticizes the use of this measure, suggesting it is not mountains *per se*, but the type of terrain that is important. This does not, however, invalidate its use here.

6. Data for the period 1960 to 1999 came from the World Bank and was cross-referenced with Fearon (2005) for consistency, and export data (primary commodities) came from the World Trade Organization (WTO) and was combined with GDP from the World Bank for the remaining years.

7. Oil *exports* are coded as a binary variable: 1 if the share of oil exports in total exports is greater than one-third (33.3%) and 0 otherwise. Oil *production* data, in metric tons annually, are provided by Ross (2013) for the years 1932 to 2011. The additional two years were drawn from Ross’ source, the U.S. Department of Energy site for international energy statistics: <http://www.eia.gov/cfapps/ipdbproject/IEDindex3.cfm>.

8. Initially used by Easterly and Levine (1997), the fractionalization index follows Herfindahl’s formula, and is interpreted as the probability that two randomly selected individuals in a population belong to different ethnic groups.

9. They argued that economic grievance is captured by the relative gap between the mean national income and the income level of the poorest and richest groups (positive and negative horizontal inequality), while ethnopolitical grievance is measured by the demographic size of the largest ethnic group discriminated against. The units of measurement are as follows: LDG = demographic size of the largest discriminated against ethnic group relative to the joint size of the discriminated group and the group in power (bound between 0 and 1); NHI = mean country GDP per capita / mean per capita income of poorest group; PHI = mean per capita income of richest group / mean country GDP per capita.

10. This differs from the existing literature, but in light of the empirical setup for Table 2, it makes some sense to find that primary commodity exports, as a share of GDP, are lower for countries not in civil conflict.

11. For a full explanation of the largest discriminated against ethnic growth (LDG), see Buhaug, Cederman, and Gleditsch (2014).

12. See Lambert (1992) and Hall (2000) for a full derivation of the model.

13. Statistical inference becomes increasingly difficult as the proportion of zeroes gets close to one.

14. To reiterate, specification (2) is merely a check on whether the choice of variables in the selection/inflation equation in specification (1) has a drastic impact on the type of results one obtains. Given that all variables are in both equations in specification (2), the results for the two outcome equations are surprisingly similar. The only noticeable differences between the two specifications are that in specification (2), the mountain variable becomes insignificant, two of the grievances terms become insignificant, and population changes sign. The

selection/inflation equation shows the probability of nonparticipation. Coefficient magnitudes can only be interpreted by calculating the marginal effects, not directly from the coefficients. For example, the variable, log real GDP has a coefficient of -0.375 , but computing the marginal effect shows that higher GDP reduces the probability of being in the “experienced conflict” group by 9.1 percent.

15. Reported in an Appendix in Dunne and Tian (2017).

16. Normally the threshold is 0.5, so a dichotomous conflict variable is equal to 1 if the estimated probability is greater than 0.5 and 0 otherwise. This is then compared to the actual. The ROC method varies the threshold between 0 and 1, creating a curve plotting the true positive rate against the false positive one. Similar to the well-known Gini coefficient procedure, the area under the ROC curve summarizes a model’s overall predictive power.

17. Since ROC's cannot be performed on variables that are not binary, the ordered outcome dependent variable (0,1,2) was divided into two binary (0,1) variables, namely, minor conflict (equivalent to the original variable equaling 1) and major conflict (equivalent to the original variable equaling 2). Separate ROC tests were then conducted to test the predictive power of the models and the individual variables on correctly predicting each type of conflict. Note that Fearon and Laitin (2003) and Collier and Hoeffler (2004) get ROC values of 0.761 and 0.860, respectively for their models. (These ROC values are taken from Ward, Greenhill and Bakke, 2010, who only ran 1 ROC each for F&L and C&H.)

18. Much the same can be said for most of the opportunity and the grievance variables but the Polity IV is of interest because while it became *statistically* significant once we switched from the probit to the ZiOP model, the Δ ROC suggests that it is not *substantively* significant.

References

- Bagozzi, B.E., D.W. Hill, W.H. Moore, and B. Mukherjee. 2015. “Modelling Two Types of Peace: The Zero-Inflated Order Probit (ZiOP) Model in Conflict Research.” *Journal of Conflict Resolution*. Vol. 59, No. 4, pp. 728–752. <https://doi.org/10.1177/0022002713520530>
- Besley, T., and T. Persson, 2010. “State Capacity, Conflict and Development.” *Econometrica*. Vol. 78, No. 1, pp. 1–34. <https://doi.org/10.3982/ECTA8073>
- Besley, T., and T. Persson. 2014. “The Causes and Consequences of Development Clusters: State Capacity, Peace and Income.” *Annual Review of Economics*. Vol. 6, pp. 927–949. <https://doi.org/10.1146/annurev-economics-080213-041128>
- Blattman, C., and E. Miguel. 2010. “Civil War.” *Journal of Economic Literature*. Vol. 48, No. 1, pp. 3–57. <https://doi.org/10.1257/jel.48.1.3>
- Boulding, K.E. 1978. *Stable Peace*. Austin, TX: University of Texas Press.
- Brauer, J., and J.P. Dunne. 2012. *Peace Economics: A Macroeconomic Primer for Violence-Afflicted States*. Washington, D.C.: United States Institute of Peace.
- Buhaug, H., L.-E. Cederman, and K.S. Gleditsch. 2014. “Square Pegs in Round Holes: Inequalities, Grievances, and Civil War.” *International Studies Quarterly*. Vol. 58, No. 2, pp. 418–431. <https://doi.org/10.1111/isqu.12068>
- Burke, M.B., E. Miguel, S. Satyanath, J.A. Dykema, and D.B. Lobell. 2009. “Warming Increases the Risk of Civil War in Africa.” *Proceedings of the National Academy of Sciences*. Vol. 106, No. 49, pp. 20670–20674. <https://doi.org/10.1073/pnas.0907998106>
- Cameron, A.C., and P.K. Trivedi. 2010. *Microeconometrics Using Stata*. Rev Ed. College Station, TX: Stata Press.
- Collier, P., and A. Hoeffler. 2004. “Greed and Grievance in Civil War.” *Oxford Economic Papers*. Vol. 56, No. 4, pp. 563–595. <https://doi.org/10.1093/oep/gpf064>
- Collier, P., and A. Hoeffler. 2007. “Civil War,” pp. 711–739 in K. Hartley and T. Sander, eds. *Handbook of Defense Economics*. Vol. 2. Amsterdam: Elsevier. [https://doi.org/10.1016/S1574-0013\(06\)02023-0](https://doi.org/10.1016/S1574-0013(06)02023-0)
- Collier, P., V.L. Elliot, H. Hegre, A. Hoeffler, M. Reynal-Querol, and N. Sambanis. 2003. *Breaking the Conflict Trap: Civil War and Development Policy*. Washington, D.C.: The World Bank. <https://doi.org/10.1037/e504012013-001>
- Dunne, J.P., and F. Coulomb. 2008. “Peace, War and International Security: Economic Theories,” pp. 13–26 in J. Fontanel and M. Chatterji, eds. *War, Peace and Security*. Vol. 6. Bingley, UK: Emerald. [https://doi.org/10.1016/S1572-8323\(08\)06002-5](https://doi.org/10.1016/S1572-8323(08)06002-5)
- Dunne, J.P., and N. Tian. 2014. “Conflict Spillovers and Growth in Africa.” *Peace Economics, Peace Science and Public Policy*. Vol. 20, No. 4. <https://doi.org/10.1515/peps-2014-0038>
- Dunne, J.P., and N. Tian. 2017. “Conflict and Fragile States in Africa.” African Development Bank. Working Paper No. 274. <https://www.afdb.org/en/documents/document/working-paper-274-conflict-and-fragile-states-in-africa-97150>.
- Easterly, W., and Levine, P. 1997. “Africa's Growth Tragedy: Policies and Ethnic Divisions.” *The Quarterly Journal of Economics*. Vol. 112, No. 4, pp. 1203–1250. <https://doi.org/10.1162/003355300555466>
- Fearon, J.D. 2005. “Primary Commodity Exports and Civil War.” *Journal of Conflict Resolution*. Vol. 49, No. 4, pp. 483–507. <https://doi.org/10.1177/0022002705277544>
- Fearon, J.D., and D.D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review*. Vol. 97, No. 1, pp. 75–90. <https://doi.org/10.1017/S0003055403000534>
- Fearon, J.D., K. Kasara, and D.D. Laitin. 2007. “Ethnic Minority Rule and Civil War Onset.” *American Political Science Review*. Vol. 101, No. 1, pp. 187–193. <https://doi.org/10.1017/S0003055407070219>

Grossman, H.I. 1991. "A General Equilibrium Model of Insurrections." *American Economic Review*. Vol. 81, No. 4, pp. 912–921.

Hall, D.B. 2000. "Zero-inflated Poisson and Binomial Regression with Random Effects: A Case Study." *Biometrics*. Vol. 56, No. 4, pp. 1030–1039.
<https://doi.org/10.1111/j.0006-341X.2000.01030.x>

Harris, M.N., and X. Zhao. 2007. "A Zero-inflated Ordered Probit Model with an Application to Modelling Tobacco Consumption." *Journal of Econometrics*. Vol. 141, No. 2, pp. 1073–1099.
<https://doi.org/10.1016/j.jeconom.2007.01.002>

Hegre, H., T. Ellingsen, S. Gates, and N.P. Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992." *American Political Science Review*. Vol. 95, No. 1, pp. 33–48.

Hendrix, C.S., and S.M. Glaser. 2007. "Trends and Triggers: Climate, Climate Change and Civil Conflict in Sub-Saharan Africa." *Political Geography*. Vol. 26, pp. 695–715.
<https://doi.org/10.1016/j.polgeo.2007.06.006>

Hirschleifer, J. 1995. "Theorizing About Conflict," pp. 165–189 in K. Hartley and T. Sandler, eds. *Handbook of Defense Economics*. Vol. 1. Amsterdam: Elsevier.
[https://doi.org/10.1016/S1574-0013\(05\)80009-2](https://doi.org/10.1016/S1574-0013(05)80009-2)

Lambert, D. 1992. "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing." *Technometrics*. Vol. 34, No. 1, pp. 1–14.
<https://doi.org/10.2307/1269547>

Lujala, P., N.P. Gleditsch, and E. Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Conflict Resolution*. Vol. 49, No. 4, pp. 538–562.
<https://doi.org/10.1177/0022002705277548>

Miguel, E., S. Satyanath, and E. Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy*. Vol. 112, No. 4, pp. 725–753.
<https://doi.org/10.1086/421174>

Pickering, S. 2011. "Determinism in the Mountains: The Ongoing Belief in the Bellicosity of 'Mountain People'." *The Economics of Peace and Security Journal*. Vol. 6, No. 2, pp. 21–25.
<https://doi.org/10.15355/epsj.6.2.21>

Ross, M.L. 2013. "Oil and Gas Data, 1932–2011." Harvard Dataverse, v.2, DOI_1902.1_20369.xml.

Salehyan, I., and K.S. Gleditsch. 2006. "Refugees and the Spread of Civil War." *International Organization*. Vol. 60, No. 2, pp. 335–366.
<https://doi.org/10.1017/S0020818306060103>

Sambanis, N. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil War." *Defence and Peace Economics*. Vol. 13, No. 3, pp. 215–243.
<https://doi.org/10.1080/10242690210976>

Skaperdas, S. 2008. "An Economic Approach to Analyzing Civil Wars." *Economics of Governance*. Vol. 9, No. 1, pp. 5–44.

<https://doi.org/10.1007/s10101-007-0043-2>

Vance, C., and N. Ritter. 2014. "Is Peace a Missing Value or a Zero? On Selection Models in Political Science." *Journal of Peace Research*. Vol. 51, No. 4, pp. 528–540.
<https://doi.org/10.1177/0022343314528200>

Ward, M.D., B.D. Greenhill, and K.M. Bakke. 2010. "The Perils of Policy by p-value: Predicting Civil Conflicts." *Journal of Peace Research*. Vol. 47, No. 4, pp. 363–375.
<https://doi.org/10.1177/0022343309356491>

Wucherpfennig, J., N.B. Weidmann, L. Girardin, L.-E. Cederman, and A. Wimmer. 2011. "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." *Conflict Management and Peace Science*. Vol. 28, No. 5, pp. 423–437.
<https://doi.org/10.1177/0738894210393217>

Appendix

Zero-inflated models

A zero-inflated ordered probit (ZiOP) model follows a two stage estimation process. The first is a selection or *inflation equation*, and the second stage is a probit *outcome equation*. This splits the observations into two processes, each potentially having different sets of explanatory variables. In the context of civil war prevalence, zero observations in process 0 ($w_i=0$) include inflated zeroes, consistent with countries that never experience civil conflict (e.g., Botswana), while zero observations in process 1 ($w_i=1$) includes cases for which the probability of transitioning into a civil conflict is not zero, but civil war casualties have not reached the lower bound of 1,000 battle-related deaths. The binary variable w indicates the split between process 0 and process 1 and is related to the latent dependent variable w_i^* , so that $w_i=1$ for $w_i^*>0$ and $w_i=0$ for $w_i^*\leq 0$, where w_i^* now represents the propensity to enter process 1, given by the split probit 1st stage or *inflation equation*:

$$(1) w_i^* = x_i\gamma + \mu_i.$$

Here, x_i is a vector of covariates, γ is its coefficients, and μ_i is the error term. The probability of country i falling into process 1 (that is, war) is $\Pr(w_i=1|x_i) = \Pr(w_i^*>0|x_i) = \mathcal{P}(x_i\gamma)$, and for process 0 (peace) it is $\Pr(w_i=0|x_i) = \Pr(w_i^*\leq 0|x_i) = 1 - \mathcal{P}(x_i\gamma)$, where $\mathcal{P}(\cdot)$ is the standard normal cumulative distribution function. For the probit 2nd stage, or *outcome equation*, the propensity for participation in which the response variable Y_i (i.e, conflict) has a distribution given by:

$$(2) \Pr(Y_i=y_i) = \begin{cases} w_i + (1-w_i)e^{-\lambda_i} & , y_i = 0 \\ (1-w_i)e^{-\lambda_i} \frac{\lambda_i^{y_i}}{y_i!} & , y_i > 0 \end{cases}$$

where the parameters λ_i and w_i depend on vectors of covariates x_i and z_i , respectively, which are modeled as $\log(\lambda_i) = x_i'\beta$ and $\log[w_i/(1-w_i)] = z_i'\gamma$, with mean and variance as $E(Y_i) = (1-w_i)\lambda_i$ and $\text{var}(Y_i) = \mu + [w_i/(1-w_i)]\mu^2$.

In this ZiOP model, the matrices z and x contain different

sets of experimental factor and covariate effects that relate to the probability of the zero-state (zero probability of civil war) and the Poisson mean in the nonzero-state (probable civil war), respectively. Thus, the γ 's have interpretations in terms of the factor level effect on the probability that there is a zero probability of conflict and the β 's have the interpretation of the effect on the average risk of civil war when the probability is nonzero. Following Lambert (1992), equation (2) in the ZiOP model can then be regressed using maximum likelihood with an expectation-maximum (EM) algorithm. For the full derivation, see Lambert (1992) and Hall (2000).