Southern Africa Labour and Development Research Unit

Determinants of Civil War and Excess Zeroes

by
John Paul Dunne and Nan Tian

Working Paper Series
Number 191

About the Author(s) and Acknowledgments

John Paul Dunne: School of Economics and SALDRU, University of Cape Town, Cape Town, 7701;
e-mail: john.dunne@uct.ac.za

Nan Tian: School of Economics and SALDRU, University of Cape Town, Cape Town, 7701;
email: nan.tian.eco@gmail.com

Recommended citation

Dunne, JP., Tian, N., (2016). Determinants of Civil War and Excess Zeroes. A Southern Africa Labour and Development Research Unit Working Paper Number 191. Cape Town: SALDRU, University of Cape Town

---

ISBN: 978-1-928281-52-8

Determinants of civil war and excess zeroes

John Paul Dunne and Nan Tian

SALDRU Working Paper Number 191
University of Cape Town
September 2016

Abstract

This paper considers the determinants of civil conflict, using a zero-inflated modelling approach that deals with the problem of excess zero observations, which we argue are related to two distinct data generation processes. Despite their continued use in the literature, traditional probit and logit models have limited capacity in dealing with this issue and can create misleading results. This is illustrated by estimating the model in Elbadawi and Sambanis (2002) using their data and a zero-inflated modelling procedure, which leads to results that suggest a role for the grievance variables in contrast to the original article. A general greed-grievance model is then estimated on a sample of 134 countries, over 54 years. Again, while the standard probit model results tend to emphasise opportunity variables, as found in other studies, the zero-inflated model gives more support for grievance effects. In particular, polity, ethnicity and inequality are found to play a significant role in contrast to earlier studies.

Keywords: Civil war; zero-inflation; greed and grievance
JEL code: D74; C3

1 Introduction

Civil war is a major concern for developing countries, as it can have profound negative effects on development, diverting resources from productive activities to destruction, destroying market and social relations and leading to death and suffering. These effects are amplified if one considers the opportunity cost of the productive resources allocated to conflict and the loss from the damage caused (Collier et al., 2003). Violent internal conflicts have ravaged many countries in the world, since 1960, the average prevalence of civil war was about 12%, with a peak of 17% in the years 1990 and 1991 (Lacina and Gleditsch, 2005; Themner and Wallensteen, 2014). This led to a concern as to why there was so much civil war in the world and an extensive literature was developed that sought to answer this question. In particular, quantitative research on the causes of civil war was undertaken by Collier and Hoeffler (1998) and Keen (1998) and Collier and Hoeffler (2004) and Fearon and Laitin (2003) and provided valuable but controversial insights into the role economic and socio-political factors in civil war initiation, continuation and potential end (Blattman and Miguel, 2010).

Most of empirical studies on the determinants of civil war use some form of probit or logit model to estimate a zero-one dependent variable, with zero for peace and one for conflict, on a panel or cross-section of countries. This has served research well, but there is a potential problem that has not been recognised. These models do not perform well in situations where there are a large number of zeroes in the dependent variable, which is likely to be the case for civil conflict, since many country-year observations are, fortunately, zero (i.e. peace). Until recently there had been little recognition of this issue, possibly because the earlier models were estimated on cross section data, five year average panels or lacked the appropriate empirical tools, but the use of annual data in panels has made it an issue of greater concern. Using a probit or logit model in the presence of excess zeroes can potentially lead to biased estimates due to the correlation of the error term with the explanatory variables (Bagozzi et al., 2015).

In such studies the value of one is recorded when the number of battle-related deaths exceeds a particular bound (e.g. 1000 or 25) and is zero otherwise. This zero can, however, reflect rather different situations. It encompasses situations when the number of battle related deaths is zero and there is complete peace and situations where there are minor conflicts but battle related deaths are still below the defined threshold. More importantly, it also includes countries with heterogeneous experiences, where for some the zero may represent a break between conflicts in a conflict ridden country, while for others it may represent a year of peace in a completely peaceful country. Consider the difference between a zero for a country such as Sweden or Australia which has almost no chance of civil war, and a zero for countries like the Democratic Republic of Congo (DRC) or Cambodia, where civil war risk is considerably higher. so the zeroes from the different countries come from different processes and not taking this into account can lead to statistically biased coefficent estimates.

Another issue is that when the conflict dependent variable has excess peace observations or zeroes, the probit and logit models cannot statistically account for the observable and latent factors that generate these high proportion of zeroes. They generate only one latent equation and are unable to account for or differentiate between the different additional weights put on zero observations, especially if the zeroes relate to different processes. Using probits or logits in hypothesis testing could lead to model misspecification (Harris and Zhao, 2007).

A further concern is whether the probit or logit models conform to the process that generated the data. Civil conflict can be characterised as an event of rare occurrence and research has suggested a Poisson distribution would better suit the data than logistical or normal distributions (Smith and Tasiran, 2012).

This paper considers the determinants of conflict, taking these issues into account and using a zero-inflated (ZIP) model. To get some idea of the likely effect of using normal logit or probit models, when a split population or zero-inflated model might be more suitable, it uses the data set of a published

study, Elbadawi and Sambanis (2002) and considers the impact on the results. A zero-inflated model is then used to revisit the greed and grievance debate with a data set of 134 countries for the period 1960 to 2013. The next section gives a brief review of the determinants of civil war literature. This is followed by an overview of the zero-inflated Poisson (ZIP) model. Section four then provides the results of the replication analysis of Elbadawi and Sambanis (2002). Thereafter, section five presents the empirical re-estimation of the greed-grievance argument, applying the ZIP model to the new data set. The final section offers some conclusions, with discussion on the implications for future research.

2 The Causes of Civil Conflict

In the studies of the determinants of civil war, the initial debate was between those who saw civil conflict as the result of grievance, a genuine or perceived injustice that needs to be redressed, and those who considered it the result of greed (Sambanis, 2002). Collier and Hoeffler (2004) formalised this in a rational choice model, where income is achieved during a rebellion from looting or, after a successful rebellion, from control of state revenues. In this case, the probability of civil conflict increases as benefits rise relative to costs. Alternatively, conflicts could be caused by differences in religion, ethnicity, income and class. A number of papers, including Fearon and Laitin (2003), Collier and Hoeffler (2004) and Elbadawi and Sambanis (2002) have extended the analysis, turning statements of correlation into causal relationships and debating the form of the variables and structure of the relationship. This empirical literature has been dominated by cross-country regressions, with a consensus developing on the factors that make countries more prone to civil war. They are considered more likely to occur in countries that are poor, have low growth rates, are easily subjected to adverse income shocks and have high dependence on primary commodity exports (Blattman and Miguel, 2010).

These causal relationships depicted by income, commonly grouped as "greed" variables, are interpreted as lack of opportunities which make rebel recruitment and participation easier (Collier and Hoeffler, 2004). The opportunity cost for rebels is considered lower for countries with low growth and income, while chances of gain from war is much higher. Countries with high proportion of primary commodity exports in their GDP are thought to be more prone to conflict due to natural resources acting as a source of finance for rebellion and weakening institutions (Humphreys, 2005; Robinson et al., 2006). This dominance of economic factors has been further backed up by a range of studies, including Hegre et al. (2001), Montalvo and Reynal-Querol (2005), Ross (2006), Besley and Persson (2014) and Hoeffler (2012), finding variables that proxy for factors such as political rights, income inequality, ethnicity and religion to add little explanatory power to civil conflict.

So far most of the discussion on the determinants of civil war have focused on civil war onset or incidence, but there have also been work on the prevalence of civil war (Elbadawi and Sambanis, 2002; Reynal-Querol, 2002; Besley and Persson, 2010). Although studies on civil war incidence consider onset, prevalence and recurrence, they share common empirical results and economic factors are generally found to most strongly predict whether civil wars occur than grievance factors..[1]

The issue of causal identification has been a prominent feature in the literature. Researchers had tried to address the potential endogeneity of conflict in a growth equation, by lagging the explanatory variables, but a more satisfactory approach has come from the use of instrumental variables. Miguel et al. (2004) use exogenous variation in rainfall as an instrument for income growth, for developing economies that still depend on agriculture and find that a 5% drop in income growth leads to an increase in civil conflict likelihood in the following year by almost 10%. This is a result that has since been supported by a number of quantitative studies, with warmer temperatures

[1]Prevalence can be defined as the likelihood of observing a civil war at any point in time $Y(t)$, estimating the probability of that $Y(t) = 1$ is the sum of the probability that war occurs at time t contingent on there being no war at $t-1$ and the probability that war occurs at time t given that war had been ongoing at time $t-1$.

or extremes in rainfall causally associated with changes in the probability of civil war onset (Hsiang et al., 2013).

Researchers have also concerned themselves with measurement and other such empirical issues. For instance, Ross (2006) argued that compared to older natural resource measures, improved measures of oil and diamond deposits are strongly associated with more civil conflict. Likewise, the lack of significance of the objective grievance variables has been strongly contested by political scientists. Most recently by Buhaug et al. (2014), who suggests that the lack of significance had to do with the poorly measured proxy variables (e.g. Gini coefficient and ethno-linguistic fractionalisation) used in previous research. By using variables that better reflect ethnic and income inequality, the authors find that political and economic grievances do matter. There is also currently little agreement on the correct econometric specification, authors vary in the use to annual versus five-year periods, the definition of civil conflict, the appropriate estimator for these rare events and the degree of measurement error in the dependent and independent variables (Blattman and Miguel, 2010). Sambanis (2004) found major differences in the datasets used by various authors, with most attributed to the different definitions of civil war, while others such as Vance and Ritter (2014) and Smith and Tasiran (2012) have questioned the estimators used to study these "rare" events.

Interestingly, looking back at the literature, one sees that almost all empirical research on civil war has used some form of probit or logit model to estimate a set of explanatory variables on a zero-one dependent variable. Yet recent analysis have suggested that research using these models can potentially run into econometric challenges due to two important issues. Firstly, the use of a zero-one indicator variable to denote conflict or peace often runs into the problem of excess number of zero observations (e.g. peace observations) These zeroes, often in the region of 80% of total observations, have the potential to be heterogeneous and come from different data generating processes. In this case, estimates that use the common probit or logit models to determine civil war risk may run into issues of model misspecification and

biased estimates (Bagozzi et al., 2015). To deal with this issue, Bagozzi et al. (2015) suggest that an alternative method, specifically a split-population or zero-inflated model, should be used to address excess zeroes in the dependent conflict variable. Using this method, they find that when the dependent variable contains excess zeroes, estimates which use zero-inflated or split population models are more reliable and consistent, compared to those from ordinary probit or logit models. Furthermore, they indicate zero-inflated models produce more accurate coefficient estimates of key independent variables on conflict outcome.

In addition, there have been concerns over whether or not probit or logit models used in estimating civil war risk conform to the process that generates the data (Smith and Tasiran, 2012). Richardson (1960) initially found civil wars starting per year to closely fit the theoretical Poisson distribution of rare events, a result confirmed by Wilkinson (1980) and Benoit (1996). It would seem that in the case of civil war, an event count process that is characterised by having a rare occurrence, a Poisson distribution is likely to better suit the data than logistical or normal distributions found when using logit or probit models respectively.

Given the present state of the literature it seem opportune to revisit the earlier debate of the determinants of civil war, the greed versus grievance debate, using more data, better variable measurement and using a more appropriate estimator for rare events.

3 Modelling Zero-Inflation in Civil Conflict

In most analyses on the determinants of civil conflict, an ordered dependent variable is used in which a given country-year is assigned a zero for peace and a value of one when violence between the state and insurgents reaches a given threshold, classifying it as a civil war. This would generally mean that there are a large number of zero observations since peaceful years will dominate conflict years. These zeros can be considered as reflecting rather

different states, one where the structural and societal forces ensure a zero probability of civil conflict regardless of greed or grievance incentives and another that reflects a break in fighting and a high probability of returning to conflict.

The first group of zeros will often be advanced or welfare economies, such as Norway, Sweden, or Japan and can be labelled "complete-peace" while the second group, are often found in developing regions such as sub-Saharan Africa, Asia or Latin America, from which the zeroes can be labelled as "incomplete-peace". The main difference between the first and second type of zero is that while the probability of transition into war for the first type is zero, the probability for the "incomplete-peace" group is not. In the case of a "incomplete-peace" incentives resulting from opportunity or grievance can induce violent conflict. There is also a third type of zero observation where there is still fighting taking place, but the violence has not reached the 1 000 battle related deaths threshold used to define a conflict. Such cases can be labelled as "incomplete-war". Given the high proportion of heterogeneous zeroes in the analysis, using ordinary probit or logit models may not be an appropriate tool for statistical inference and can potentially give biased estimates (Bagozzi et al., 2015).

To shed some light on the above mentioned types of zeroes, the world map in Figure 1 below provides a breakdown of countries that are categorised as "complete peace" (white), "incomplete-peace or incomplete war" (light grey) and "complete-war"(dark grey).[2] Interpretation of the illustrated map is intuitive, countries in the least developed regions of the world, such as Africa, South-east and Central Asia are often found to be in the category of "incomplete-peace" or "complete-war". By contrast, highly developed countries and democracies, such as those found in Europe, North America

[2]Countries highlighted in white are those that have never experienced a civil conflict and thus labelled as "complete peace" . Those coloured in light grey are countries that have experienced some form of minor civil violence (e.g. between 25 and 999 battle related deaths in a given year) over the past 54 years and can be considered as "incomplete-peace" and "complete-war" . Countries in dark grey are those that have had intense civil conflicts totalling over 1 000 battle related deaths in a year and can be considered as a truly conflict group.

and parts of Oceania almost always belong to the "complete peace" group.

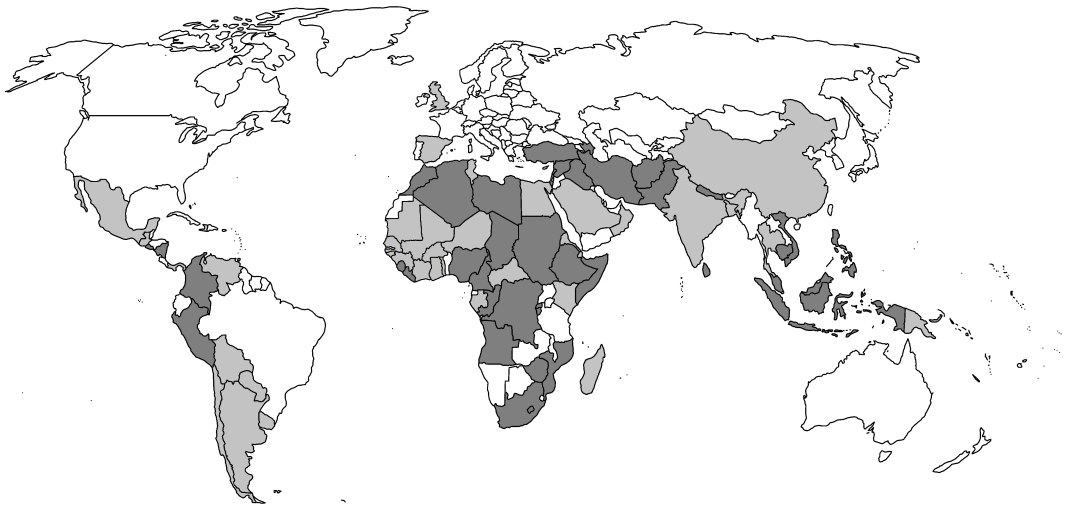

Figure 1: Complete peace, incomplete peace and civil war intensity

With such a clear difference in the meaning of zeroes, a more satisfactory estimation alternative is to use a split population or two-part model as proposed by Harris and Zhao (2007) and Vance and Ritter (2014). This is typically in the form of zero-inflated models, or in this case a zero-inflated Poisson model, where estimations follow two stages. The first of the two latent equations, stage one, is a selection equation, while the second stage is a Poisson 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 that never experience civil conflict (e.g. Sweden), while zero observations in process 1 ($w_i = 1$) includes cases for which the probability of transitioning into a civil conflict is not zero and civil war casualties have not reached the lower bound (or limit) of 1 000 battle related deaths. The binary variable w indicates the split between process 0 (with $w_i = 0$ for no war) and process 1 (with $w_i = 1$ for war). w 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 and is given by the split probit (1st stage) equation:

$$w_i^* = x_i\gamma + \mu_i \tag{1}$$

where x_i is a vector of covariates, γ is its coefficients and μ_i is the error term. The probability of i falling into process 1 is $\Pr(w_i = 1|x_i) = \Pr(w_i^* > 0|x_i) = \Psi(x_i g)$, and the probability that it is in process 0 is $\Pr(w_i = 0|x_i) = \Pr(w_i^* <= 0|x_i) = 1 - \Psi(x_i\gamma)$, where $\Psi(.)$ is the standard normal cumulative distribution function. For the Poisson outcome equation, the propensity for participation in which the response variable Y_i (i.e conflict) has a distribution given by:

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

where the parameters λ_i and w_i depend on vectors of covariates x_i and z_i, respectively, which are modelled as:

$$\log(\lambda_i) = x_i^{'}\beta \tag{3}$$

and

$$\log(\frac{w_i}{1 - w_i}) = z_i^{'}\gamma \tag{4}$$

with mean and variance as $E(Y_i) = (1-w_i)\lambda_i$ and $var(Y_i) = \mu + \left(\frac{w_i}{1-w_i}\right)\mu^2$. In this ZIP 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 non-zero. Following Lambert (1992)

the ZIP model (equation 2) can be estimated using maximum likelihood with an Expectation-maximum (EM) algorithm.[3]

While empirical research on the determinants of civil conflict has generally followed the standard approach of assuming normality, conflict data is produced in a discrete and countable manner with the number of events never being negative or a non-integer. This does not suggest a normal or logistical distribution and the error terms in a regression would not be normally distributed and the observed variables would not be a linear function of the covariates (Benoit, 1996). When civil conflict is a random event, and is observed at the end of each observation period i (common in conflict studies), then the data will conform to a Poisson process which has a rate of occurrence λ, where $\lambda > 0$, as long as the event occurred at the start of the period and no more than one event occurs at the same time. Both assumptions are satisfied in civil war research, as firstly, by definition, no more than one civil war can occur in a given country. Secondly, since conflicts are only recognised if the threshold (either more than 25 or 1000 battle-related deaths) has been reached at the end of a calendar year, occurrences in the previous period are considered independent events. In other words, conflict events are seen to be both *stationary* and *independent*. Additionally, Benoit (1996) stated the possibility of over-dispersion in the Poisson estimates and suggested using negative binomial as an alternative. Analysis of mean and variance shows no significant difference between the two, thus using the Poisson distribution seems theoretically more suitable and has been shown to fit the distribution of conflicts over time.

The use of ZIP model allows more accurate estimates to be obtained compared to standard probit or logit models. The probability of a zero observation is now modelled conditional on the probability of zero from the Poisson process plus the probability of being in process 0 from the splitting equation. 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

[3]For full derivation of the model see Lambert (1992) and Hall (2000).

or very small, leading to biased results.[4] Bagozzi et al. (2015) suggests that this will become an issue when there is less than 10 percent or greater than 90 percent of zero observations.

4 An Empirical Investigation

Elbadawi and Sambanis (2002) provided an influential contribution to the "greed-grievance" debate by combining Collier and Hoeffler (2004) model of civil war onset with the Collier et al. (2004) model of civil war duration. Their model predicts the prevalence of civil conflict, based on opportunities for rebellion against its constraints, with opportunities divided into greed versus grievance, or rebellions that generate profit versus rebellions triggered by genuine grievance. They code incidents as civil conflicts using five categories, those that caused more than 1 000 battle-related deaths, those that challenged the sovereignty of an internationally recognised state, those that occurred in the territory of the state, those that included the state as a principle combatant, and those where the rebels were able to mount a organised military opposition to the state. Their sample includes over 150 countries between the period 1960 to 1999, with a dependent variable having about 81 percent of zero or peace observations. Using a probit model, Elbadawi and Sambanis find that prevalence of civil war is consistent with earlier studies on war onset and duration (Table 1, column 1). It is positively influenced by primary commodity exports as a share of GDP (a proxy for "looting" or economic opportunity), population and previous wars experiences in the past 10 years, while the level of GDP, the growth of per capita GDP and squared term of primary commodity exports as a share of GDP have negative effects. As with most studies in the literature, grievance factors such as ethnic fractionalisation, ethnic dominance and polity (proxy for political rights) were statistically insignificant.

Given the potential for high proportion of heterogeneous zeroes, which can

[4]Statistical inference becomes increasingly difficult as the proportional of zeroes gets close to one.

be thought of as including "complete-peace" or "incomplete peace" a zero-inflated Poisson model (ZIP) would seem more suitable that the usual probit model. Table 1 presents the results for the Elbadawi and Sambanis (2002) specification and data, using a probit specification to replicate their results (1) and the ZIP model for comparison (2, with column 3 giving the first stage (Outcome) and the last column the second stage (inflation) estimates. To remain consistent with Elbadawi and Sambanis, the covariates in the outcome equation of the ZIP model are identical to the normal probit. The variables that are included in the inflation are those that are considered to directly influence the probability that a country in any given year always experiences peace (i.e. per capita GDP, ethnic diversity or political freedom).

The results for the ZIP inflation equation, show real GDP to have a positive and significant effect on the likelihood that a country-year is among among the always zero or "complete peace" group, as expected.[5] Interestingly, primary commodity exports as a share of GDP has no significant effect, with ethnolinguistic diversity having a significant effect, which is estimated to be parabolic, first decreasing and then increasing. In the inflation equation, the coefficients represent the factor change in the probability of being completely peaceful compared to incomplete peace. When the same variables are included in both the outcome and inflation equations the signs of the corresponding coefficients from the inflation equation are often opposite in sign to those in the outcome equation. This makes sense since the inflation equation is predicting membership in the group that is always peaceful, while the outcome equation predicts conflict risk conditional on a country being in a not always peaceful group.

[5]Thus, the higher a country's real GDP, the greater chance of it being always peaceful. In numerical terms, if a country increases its real GDP by 1%, the likelihood of the country being in the "complete peace" group would increase by a factor of 1.001 (e.g. exp(0.0018)).

Table 1: Probit and ZIP of Civil War Prevalence 1960-1999

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Pri Exports/GDP	10.53*	9.488*	-2.276
	(4.136)	(4.341)	(5.061)
Pri Exports/GDP^2	-21.24*	-23.41*	
	(9.325)	(11.31)	
log real GDP	-0.0003**	-0.0004**	0.0018*
	(0.000)	(0.000)	(0.0009)
RGDPPC Growth	-0.0899**	-0.122**	
	(0.0286)	(0.034)	
Polity Index (1 lag)	-0.0115	0.0135	
	(0.020)	(0.0184)	
Polity Index^2 (1 lag)	0.0032	0.0035	
	(0.0041)	(0.0042)	
Ethno Diversity	0.0389	0.0656**	-0.231**
	(0.0258)	(0.0178)	(0.0759)
Ethno Diversity^2	-0.0004	-0.0007**	0.0024**
	(0.0003)	(0.0002)	(0.0008)
Log Population	0.599**	0.266*	-2.532**
	(0.140)	(0.107)	(0.946)
Constant	-12.48**	-6.805**	
	(2.569)	(1.962)	
Rho	0.601**	-	
	(0.086)	-	
Observations	783	783	
Zero Observations	-	692	
Log likelihood	-189.771	-223.663	

Continued on next page

Table 1 – *Continued from previous page*

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Wald χ^2	33.11	-	
Vuong test	-	3.41	
AIC	479.33	401.54	

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

Both the probit and and the ZIP estimates for the outcome equation (1) and (2) show similar coefficients for most of the variables. Primary commodity exports as a share of GDP is highly significant and has a non-linear effect in predicting the probability of civil conflict, GDP levels and growth per capita, are negative and significantly related to civil war risk. There are, however, interesting difference in the grievance terms. While the Polity index remains insignificant, ethnolinguistic diversity become significant in the ZIP mode, having a non-linear effect of first increasing and then decreasing the risk of civil war prevalence. Vuong (1989) tests and Akaike Information Criterion (AIC) both favoured the zero-inflated Poisson model over the traditional probit estimator. As a robustness check across the different specifications used by Elbadawi and Sambanis, ethnic diversity was replaced with ethnic dominance and the results are presented in Table 5 in the Appendix. They provide consistent results for ethnolinguistic diversity, but with the additional polity index also becoming significant. All of the coefficients in the ZIP model, except log population and primary commodity exports as a share of GDP have larger coefficients than in the standard probit, possibly suggesting that not allowing for zero-inflation leads to civil war risk being underestimated.

These results using the Elbadawi and Sambanis (2002) data provide a strong case for moving from standard probit and logit models to some form

of a zero-inflated model in studying the determinants of conflict. Otherwise, researchers risk both underestimating the risk of civil conflict and coming to erroneous conclusions regarding its determinants. Having established this, the next section develops the analysis by estimating a more general greed-grievance model based upon the literature and using data for the period 1960 to 2013.

5 Greed vs. Grievance Revisited

To specify a general greed grievance empirical model, data for a range of variables were collected based upon the debates in the literature. Proxies for greed or opportunity include real GDP, growth in GDP per capita, degree of urbanisation, life expectancy and natural resource dependence and two sets of income variables were collected, from the World Bank and Penn World Tables 8.0. The degree of urbanisation is measured as the proportion of a country's population living in an urban environment, while life expectancy follows the usual measurement.[6] Male secondary school enrolment was not used in the estimations due to poor and incomplete data. Following the literature, natural resource dependence is measured by the share of primary commodity exports in GDP, which the World Bank provides for 1960 to 1999, and which was cross referenced with Fearon (2005) for consistency. The remaining 14 years are constructed using export data (primary commodities) provided by the World Trade Organisation (WTO) and GDP from the World Bank.

Given the numerous debates on the measure of natural resource dependence and the type of commodities used, three additional measurements were considered. Measures of oil production in metric tons and oil exports greater than one-third of total exports were used to proxy for oil abundance and

[6]This data is sourced from the World Bank, the degree of urbanisation can also be thought of as a measurement of geographic dispersion, the greater the urbanisation, the lower the geographic dispersion. All income figures are purchasing power parity (PPP) adjusted.

dependence respectively.[7] A mineral dependence variable was also created, to distinguish between fuel and non-fuel minerals and other primary commodities, with a country considered mineral dependent if its mineral exports constitute 25% or more of a its total tangible exports. The percentage of mountainous terrain in a given country was also included as an indicator of military accessibility or safe havens for rebels.

The grievance variables are, for the most part, common to those identified by Collier and Hoeffler (2004) and Fearon and Laitin (2003), with three general measures of grievance: ethnic and religious hatred, political repression or freedom, and income inequality (horizontal inequality). Ethnic fractionalisation is the most commonly chosen indicator to test the linkage between ethnicity and civil conflict.[8] Other measures taken from Collier and Hoeffler (2004), with ethnic dominance, a binary variable taking on the value one if the largest ethnic group in a country comprises between 45% - 90% of the population, and a religious fractionalisation index, which is analogous to ethnic fractionalisation.

Other things being equal, political democracy or freedom should be associated with less discrimination, repression and civil war. Data from the Polity IV database is used to measure political rights, with the variable *polity* ranging from -10 (high autocracy) to 10 (high democracy). The relationship between political freedom and civil war has often been thought of having a non-linear effect (Hegre et al., 2001) and this is tested through the inclusion of *polity* squared term. In a recent paper by Buhaug et al. (2014), the authors found that new grievance indices of horizontal income inequality and political discrimination performed much better than conventional indicators. 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

[7]Oil production in metric tons is provided by Ross (2013), this data goes from 1932 to 2011, the additional two years were drawn from the same source as the author, US energy information administration website for international energy statistics: http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm.

[8]Initially used by Easterly and Levine (1997) the fractionalisation index follows Herfindahl's formula and is interpreted as the probability tha two randomly selected individual in a population belong to different ethnic groups.

(positive and negative horizontal inequality), while ethno-political grievance is measured by demographic size of the largest discriminated ethnic group.[9] These alternative variables are used in robustness in the empirical analysis. The control variables included in the model are the standard ones found in the literature. Population, cold war and Africa feature in various specifications with their effects on civil war prevalence, apart from population, subject to much empirical debate. Finally, the dependent variable used takes on a value of 0 for all peace year observations and a 1 for civil war years with combat deaths ranging between 25-999 and annual battle deaths of above 1000.

Table 2 presents the descriptive statistics, with a breakdown by conflict experience and "complete peace" . These results seem to support the central thesis that the different zeroes in the sample are formed through completely separate processes. For the complete peace group, GDP per capita, per capita GDP growth, rate of urbanisation, life expectancy and political freedom are all higher than the conflict experience group, while countries that are potentially completely peaceful all have lower levels of ethnic and religious fractionalisation and income inequality.[10] Estimated correlations suggest some association between the income and inequality variables and the likelihood of a country being completely peaceful. In episodes of civil conflict, GDP per capita, GDP growth, rates of urbanisation, life expectancy and political freedom are all lower than in times of peace. Similarly, ethnic divisions, income inequality and substantial amounts of rough terrain are higher in cases of civil war. Interestingly, primary commodity exports as a share of GDP is on average lower in episodes of civil war compared to no civil war.

[9]For full description and derivation of the variables see Buhaug et al. (2014)

[10]ldg = largest discriminated ethnic group, phi = positive horizontal inequality (relative gap between mean national income and income level of the richest group), nhi = negative horizontal inequality (relative gap between mean national income and income level of the poorest group)

Table 2: Descriptive Statistics - Means

	Full Sample	Always 0	Not Always 0	Civil war	No Civil war
Opportunity					
Primary Commodity Exports/GDP	0.156	0.178	0.139	0.109	0.164
GDP per cap	7931	14069	3311	3172	8699
GDP per cap growth	0.018	0.022	0.016	0.010	0.019
Mountains %	16.38	14.93	18.11	23.16	15.33
Rate of Urbanisation	46.94	56.00	39.73	40.61	47.92
Life Expectancy	61.61	66.15	57.98	59.41	61.95
Oil Production (Metric Tons 000's)	17000	13700	19300	19100	16700
Mineral Dependence	0.493	0.415	0.545	0.550	0.484
Oil Exports	0.187	0.155	0.208	0.168	0.189
Grievance					
Ethnic Frac (C&H)	63.02	52.06	69.85	77.47	60.05
Ethnic Dominance	0.470	0.483	0.467	0.549	0.457
Religious Frac (0-100)	36.47	36.07	36.58	0.36	0.37
Polity IV (-10 to 10)	1.13	3.84	-0.73	0.97	1.30
LDG	0.056	0.024	0.081	0.142	0.042
NHI	1.189	1.064	1.278	1.398	1.155
PHI	1.201	1.086	1.287	1.224	1.197

Estimating the probability of civil conflict, with the dependent variable taking the value of one if deaths total over 25 in a given battle or over 1000 in a given year and zero otherwise, gave the results in Table 3. The standard probit results show GDP and per capita GDP growth to be highly significant and negative, with primary commodity exports as a share of GDP also highly significant and non-linear. Primary commodity exports are seen to initially decrease civil war risk, reaching a trough when it constitutes about 33% of

GDP, thereafter increasing civil war risk.[11] This is the opposite to that found in previous studies, but in light of the summary statistics in Table 2, where primary commodity exports as a share of GDP is lower for countries not in conflict, the result makes some empirical sense. Turning to the objective grievance terms, only the Polity IV index squared is significant with none of the measures for ethnic and religious diversity (Collier and Hoeffler (2004)'s definition) having any explanatory power. As for the control variables, population has a positive and significant effect on civil war prevalence while the Cold War dummy is negative and statistically insignificant. The likelihood ratio test of the correlation coefficient (*rho*) suggests panel estimator to be preferred to a pooled estimator. To see if a zero-inflated model would fare better in the context of excess zeroes observations, the same specification is estimated using a zero-inflated Poisson model.

Table 3: Probit and ZIP Regression of Civil War Prevalence 1960-2013

	(1)	(2)		(3)	
	Probit	ZIP		ZIP	
	Outcome	Outcome	Inflation	Outcome	Inflation
Opportunity					
log RGDP	-0.129*	-0.988**	2.551**	-0.962**	1.129**
	(0.063)	(0.032)	(0.386)	(0.038)	(0.110)
RGDPPC Growth	-2.396**	-2.697**	9.793†	-2.242**	2.392*
	(0.506)	(0.432)	(5.412)	(0.473)	(1.059)
Pri Exports/GDP	-4.145**	-5.243**		-2.842**	10.377**
	(1.107)	(0.747)		(0.819)	(2.231)
Pri Exports/GDP^2	6.593**	7.324**		3.177*	-19.618**
	(1.571)	(1.240)		(1.257)	(3.420)
log % Mountains	0.020	0.057*		0.060†	0.077
	(0.094)	(0.025)		(0.025)	(0.072)

Continued on next page

[11] Differentiate the probability of civil war with respect to primary commodity exports (4.197/2(*6.436)) = 0.326

Table 3 – *Continued from previous page*

	(1)	(2)		(3)	
	Probit	ZIP		ZIP	
	Outcome	Outcome	Inflation	Outcome	Inflation
Grievance					
Polity Index	0.004	0.017**	-0.534*	0.051*	-0.270**
	(0.007)	(0.006)	(0.218)	(0.026)	(0.074)
Polity Index^2	-0.010**	-0.009**	0.070*	-0.013**	0.003
	(0.001)	(0.001)	(0.029)	(0.004)	(0.008)
Eth Frac (C&H)	0.032	0.011†	-0.043**	0.014†	-0.084**
	(0.024)	(0.006)	(0.005)	(0.008)	(0.014)
Eth Frac^2(C&H)	-0.001	-0.001**		-0.001†	0.003**
	(0.001)	(0.000)		(0.000)	(0.001)
Ethnic Dominance	0.326	0.298**	-0.255	0.357**	-0.123
	(0.414)	(0.098)	(0.283)	(0.111)	(0.233)
Religious Frac	-0.272	-0.193		1.085**	-3.393
	(0.682)	(0.171)		(0.239)	(2.183)
Controls					
log Population	0.498**	0.112**	-3.080**	0.246**	-0.998**
	(0.125)	(0.02)	(0.444)	(0.055)	(0.131)
Cold War	0.084	0.083		0.220**	0.711**
	(0.082)	(0.067)		(0.075)	(0.191)
Constant	7.099**	-2.979**	9.950**	-6.086**	-11.339**
	(1.699)	(0.513)	(4.093)	(0.651)	(1.287)
Observations	4286	4286		4286	
Zero Observations	-	3382		3382	
Log likelihood	-1314.14	-2358.58		-2343.27	
Wald χ^2	159.34	-		-	
Vuong test	-	7.70		8.33	
AIC	4910.51	4761.18		4742.53	

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

Starting with the inflation equation, the variables of real GDP, per capita GDP growth, political freedom, ethnic diversity all represent plausible indicators that influence the probability that a country always experiences peace. To this end, the inflation equation (2) confirms that higher income and political freedom does indeed lead to greater probability of being in the "complete peace" group, while ethnic fractionalisation, ethnic dominance and population all have a negative effect. The variation in sign between the outcome and inflation estimates makes sense as one equation calculates the likelihood of countries being in conflict and the other on the probability of being in the completely peaceful group.

Moving onto the outcome equation, the zero-inflated Poisson model (2) gives signs that are consistent with the standard probit, but there are substantial differences in the significance of the grievance terms. Primary commodity exports as a proportion of GDP shows the same effect as the probit model, albeit at a higher turning point of 36%. Income per capita, both level and growth, decreases the likelihood that a country experiences civil conflict conditional on that country not being completely peaceful. Proxies for ethnopolitical grievance are better represented using the zero-inflated models than the standard probit model, with political freedom and ethnolinguistic diversity now significant predictors of civil war prevalence. Both of these variables feature in an inverse u-shape of first increasing and then decreasing civil war risk. One intriguing finding is that using the ZIP model, percentage of mountains terrain, a proxy for geographic dispersion which inhibits government or military capacity, and ethnic dominance became positive and statistically significant in explaining civil war prevalence. In addition to having more explanatory power and significance in the grievance variables, the ZIP estimates are shown to have lower standard errors and higher log likelihood values than the normal probit, suggesting a more accurate estimation outcome. The Vuong test and AIC both reject the probit model, favouring the

zero-inflated Poisson. All regressions are estimated using robust standard errors, while the proportion of zero observations in the sample (78.9%) falls within the accepted band of 10 to 90% (Bagozzi et al., 2015).

To ensure that the zero-inflated Poisson estimates are not driven by the decision to include specific variables relating to peace, (e.g. income, political freedom etc.) a full model in which all the covariates in the outcome equation included in the inflation equation. The results shown in (3), the last two columns, are consistent with the more selective specification (2). This might be considered a check for researchers' degrees of freedom, showing specifications have not been chosen to generate significant results or false positives (Simmons et al., 2011). Finding similar results for specification (2) and (3) suggests the model is well specified, consistent with various specifications and not sensitive to researcher's degrees of freedom.

To consider the robustness of the results, a number of alternative specifications are considered. Table 4 uses horizontal income inequality and ethnic discrimination in place of ethnic dominance and religious fractionalisation. The results remain consistent with earlier estimates, where the zero-inflated Poisson model is preferred to the probit model in almost all instances. Intriguingly, after changing the specification, the Cold War variable becomes negative and statistically significant, probably reflecting the decline in the number of civil wars after 1990 Blattman and Miguel (2010). Adding variables on income inequality provides a new dynamic to the civil war prevalence process. Estimating the probability of being completely peaceful, the results from the inflation equation suggest that high income inequality between the poorest group and the country average decreases the probability of peace, while increases in income inequality between the richest group and the country average increases it. [12] In the outcome equation, only negative horizontal

[12]The two horizontal income inequality measures offer some insight into the process of a country's peacefulness. Inequality has less meaning to the richest group, but is substantially more important to the poorest group. The chances of not being in "complete peace" or likelihood of violence is driven by the poorest group, which form the majority of a country's population. The richest has an incentive to keep peace as it will retain their wealth, while the poorest will have an incentive to work and overturn the inequality.

inequality (relative gap between mean national income and income level of the poorest group) has a significant positive effect on the likelihood of civil war. Higher ethnic discrimination, measured as the proportion of the largest discriminated ethnic group to the group in power, is also estimated to increase the likelihood of civil war. Similarly, the ZIP reports higher log likelihood values and smaller standard errors, with the Vuong test and AIC concluding that it is preferred to the standard probit.

Table 4: Probit and ZIP Regression of Civil War Prevalence with Income Inequality 1960-2013

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Opportunity			
log real GDP	-0.180*	-1.002**	1.115**
	(0.073)	(0.037)	(0.208)
RGDPPC Growth	-2.537**	-2.039**	2.035*
	(0.507)	(0.543)	(1.037)
Pri Exports/GDP	-4.335**	-3.783**	
	(1.100)	(0.854)	
Pri Exports/GDP^2	6.581**	4.045**	
	(1.569)	(1.453)	
log % Mountains	0.013	0.096*	
	(0.093)	(0.028)	
Grievance			
Polity Index	0.003	0.045**	-0.098*
	(0.007)	(0.007)	(0.051)
Polity Index^2	-0.010**	-0.012**	0.023**
	(0.001)	(0.001)	(0.006)
Eth Frac (C&H)	0.040†	0.026**	0.013
	(0.022)	(0.005)	(0.085)

Continued on next page

Table 4 – *Continued from previous page*

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Eth Frac^2 (C&H)	-0.001	-0.001**	
	(0.001)	(0.000)	
ldg	0.821**	1.161**	-19.76**
	(0.220)	(0.173)	(3.415)
phi	-0.137	-0.012	1.346**
	(0.150)	(0.070)	(0.383)
nhi	0.561**	0.088*	-0.969**
	(0.245)	(0.045)	(0.268)
Controls			
Population	0.552**	0.202**	-2.311**
	(0.124)	(0.045)	(0.380)
Cold War	0.008	-0.152*	
	(0.008)	(0.076)	
Constant	7.394**	-3.405**	4.450*
	(1.596)	(0.619)	(2.248)
Observations	4390	4390	
Zero Observations	-	3481	
Log likelihood	-1322.63	-1945.04	
Wald χ^2	169.66	-	
Vuong test	-	9.30	
AIC	4838.92	3940.131	

Notes: AIC = Akaike Information Criterion; Dependent variable: Conflict prevalence; Standard errors in parentheses; Significance levels:** p<0.01,* p<0.05,† p<0.1; ldg = largest discriminated ethnic group, phi = positive horizontal inequality (relative gap between mean national income and income level of the richest group), nhi = negative horizontal inequality (relative gap between mean national income and income level of the poorest group)

Other variables were considered, with primary commodity exports replaced with mineral dependence, oil production or oil export; the polity index replaced with the freedom house measure, democracy and autocracy dummies; the income variables replaced by the urbanisation rate and life expectancy; and an Africa dummy added. The results were relatively robust, with primary commodity dependence increasing civil war risk, democracy, political freedom and higher urbanisation decreasing civil war risk, and the Africa dummy having no special effect.[13]

It would appear that the models are well specified and the results are robust across a range of different indicators. This leaves us with strong support for some interesting findings. Unlike most of the literature on civil war onset, the results suggest that both greed and grievance variables are important. Using, the zero-inflated Poisson can be readily justified and allows account to be taken of observable and latent factors that produce different types of peace observations. It provides an improvement on the standard probit model and as in the repliaction of Elbadawi and Sambanis (2002) shows both ethno-political and economic grievance matter to have substantial explanatory power in predicting civil war risk.

6 Conclusion

This paper has made a contribution to the burgeoning literature on the determinants of civil conflict by highlighting a problem with using the standard probit or logit estimation procedure, when the binary conflict dependent variable is characterised by excess zeroes. In such cases, the zeroes are not homogeneous as there is a big difference between a zero that reflects peace in a peaceful country and one that reflects a lull in conflict or where the number of battle deaths used in the definition of conflict falls below the threshold used to construct the variable. Standard models are unable to account for this and a more satisfactory approach has been suggested by Bagozzi et al.

[13]See Appendix Table 6 and 7 for the additional results.

(2015), a zero-inflated model, which treats the excess zeroes as a heterogeneous group of observations, accounting for observable and unobservable factors that produce the different types of zeroes. Applying this method to the Elbadawi and Sambanis (2002) empirical specification and data led to an important difference in the results. Rather than the opportunity variable coefficients being significant, the grievance terms became significant predictors of civil war. The support of the study for the dominance of greed over grievance was not long so clearcut.

Undertaking a similar exercise on a dataset of 134 countries for the period 1960 and 2013 and using a general greed-grievance empirical specification, provided further support for the use of the zero inflated model. Using the probit model suggested that the explanation of civil war was 'greed', but using the zero inflated model suggested that both greed and grievance were important. Income, natural resource dependence, polity, ethnicity and income inequality were all found to be significant in explaining civil war prevalence. These are important results for future research and suggest that if these models had been used earlier the trajectory of the greed versus grievance debate might have been quite different.

References

Bagozzi, B. E., Hill, D. W., Moore, W. H., and Mukherjee, B. (2015). Modelling two types of peace: the zero-inflated order probit (ziop) model in conflict research. *Journal of Conflict Resolution*, 59(4):728–752.

Benoit, K. (1996). Democracies really are more pacific (in general): Reexamining regime type and war involvement. *Journal of Conflict Resolution*, 40(4):636–657.

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

Besley, T. and Persson, T. (2014). The causes and consequences of devel-

opment clusters: state capacity, peace and income. *Annual Reviews of Economics*, 6:927–949.

Blattman, C. and Miguel, E. (2010). Civil war. *Journal of Economic Literature*, 48(1):3–57.

Buhaug, H., Cederman, L., and Gleditsch, K. S. (2014). Square pegs in round holes: inequalities, grievances and civil war. *International Studies Quarterly*, 58:418–431.

Collier, P., Elliot, V., Hegre, H., Reynoal-Querol, M., and Sambanis, N. (2003). *Breaking the conflict trap.* Civil war and development policy. World Bank and Oxford University Press.

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

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

Collier, P., Hoeffler, A., and Soderbom, M. (2004). On the duration of civil war. *Journal of peace research*, 41(3):253–273.

Elbadawi, I. and Sambanis, N. (2002). How much war will we see? explaining the prevalence of civil war. *Journal of conflict resolution*, 46(3):307–334.

Fearon, J. D. (2005). Primary commodity exports and civil war. *Journal of Conflict Resolution*, 49(4):483–507.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency and civil war. *The american political science review*, 97(1):75–90.

Hall, D. B. (2000). Zero-inflated poisson and binomial regression with random effects: a case study. *Biometrics*, 56:1030–1039.

Harris, M. N. and Zhao, X. (2007). A zero-inflated ororder probit model, with an application to modelling tobacco consumption. *Journal of Econometrics*, 141(2):1073–1099.

Hegre, H., Ellingsen, T., Gates, S., and Gleditsch, N. P. (2001). Towards a democratic civil peace? democracy, political change and civil war. *American Political Science Review*, 95(1):33–48.

Hoeffler, A. (2012). *On the causes of civil war*. The oxford handbook of the economcis of peace and conflict. Oxford University Press.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, 341(6151):1235367.

Humphreys, M. (2005). Natural resources, conflict, and conflict resolution: uncovering the mechanisms. *Journal of Conflict Resolution*, 49(4):508–537.

Keen, D. (1998). The economic functions of violence in civil wars (special issue). *The Adelphi Papers*, 38(320):1–89.

Lacina, B. and Gleditsch, N. P. (2005). Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population*, 21:145–166.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14.

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

Montalvo, J. G. and Reynal-Querol, M. (2005). Ethnic polarization, potential conflict and civil wars. *American Economic Review*, 95(3):796–816.

Reynal-Querol, M. (2002). Political systems, stability and civil wars. *Defence and Peace Economics*, 13(6):465–483.

Richardson, L. F. (1960). *Statistics of deadly quarrels*. Boxwood Press, Pittsburg.

Robinson, J. A., Torvik, R., and Verdier, T. (2006). Political foundations of the resource curse. *Journal of Development Economics*, 79:447–468.

Ross, M. L. (2006). A closer look at oil, diamonds, and civil war. *Annual Review of Political Science*, 9:265–300.

Ross, M. L. (2013). Oil and gas data, 1932-2011. Harvard Dataverse.

Sambanis, N. (2004). What is civil war? conceptual and empirical complexities of an operational definition. *Journal of Conflict Resolution*, 48(6):814–858.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positives psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366.

Smith, R. and Tasiran, A. (2012). The onset of peace. *Economics of Peace and Security Journal*, 7(1):5–14.

Themner, L. and Wallensteen, P. (2014). Armed conflicts, 1946-2013. *Journal of Peace Research*, 51(4):541–554.

Vance, C. and Ritter, N. (2014). Is peace a missing value or a zero? on selection mmodel in political science. *Journal of Peace Research*, 51(4):528–540.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypothesis. *Econometrica*, 57(2):307–333.

Wilkinson, D. O. (1980). *Deadly quarrels: Lewis F. Richardson and the Statistical Study of War*. University of California Press, Berkeley, CA.

Appendix

Table 5: Probit and ZIP Models of Civil War Prevalence, 1960 to 1999 using Alternative Specification

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Pri Exports/GDP	10.566*	10.889**	
	(3.835)	(4.107)	
Pri Exports/GDP^2	-20.787*	-22.812*	
	(8.646)	(10.493)	
log real GDP	-0.0002**	-0.0002*	0.0018*
	(0.0001)	(0.0001)	(0.0006)
RGDPPC Growth	-0.0723**	-0.1461**	
	(0.0270)	(0.034)	
Polity Index (1 lag)	-0.0105	0.0341*	0.0674*
	(0.0183)	(0.0189)	(0.0301)
Polity Index^2 (1 lag)	0.0030	0.0133*	
	(0.0039)	(0.0055)	
Ethic Dominance	0.3619	0.5265*	1.0852
	(0.2909)	(0.2433)	(0.7788)
Log Population	0.4290**	0.4291**	-1.1070**
	(0.1217)	(0.1222)	(0.3131)
War in Past 10 Years	0.7353**	1.4194**	
	(0.2137)	(0.2629)	
Constant	-9.548**	-6.805**	11.753*
	(2.223)	(1.962)	(4.708)
Rho	0.4747**	-	
	(0.1049)	-	
Observations	783	783	

Continued on next page

Table 5 – *Continued from previous page*

	(1)	(2)	
	Probit	ZIP	
	Outcome	Outcome	Inflation
Zero Observations	-	692	
Log likelihood	-184.481	-211.817	
Wald χ^2	51.82	-	
Vuong test	-	3.68	

Notes: Standard errors in parentheses; Significance levels:** p<0.01,* p<0.05, † p<0.1

Table 6: Additional Results of ZIP Model: Varying in Primary Commodity

	(1)		(2)		(3)	
	ZIP		ZIP		ZIP	
	Outcome	Inflation	Outcome	Inflation	Outcome	Inflation
Opportunity						
Min Dependence	0.287*					
	(0.075)					
Oil Export			0.148			
			(0.108)			
Oil Production					-0.042	
					(0.140)	
Oil Production^2					0.002	
					(0.005)	
log real GDP	-0.092*	-0.409**	-0.070	0.416*	-0.306**	0.249*
	(0.042)	(0.064)	(0.042)	(0.061)	(0.069)	(0.114)
RGDPPC Growth	-2.079**	-2.067	-2.342**	-1.418	-3.080**	-2.343
	(0.589)	(1.837)	(0.623)	(1.198)	(0.723)	(1.809)
log % Mountains	0.127**		0.101**		0.137**	
	(0.027)		(0.026)		(0.038)	
Grievance						
Polity Index	0.043**	0.078**	0.044**	0.073**	0.080**	0.318**
	(0.008)	(0.021)	(0.008)	(0.019)	(0.009)	(0.062)
Polity Index^2	-0.004**		-0.003**		-0.004**	
	(0.001)		(0.001)		(0.002)	
Eth Frac (F&L)	5.265**	1.191**	5.367**	1.187**	5.254**	1.123
	(0.737)	(0.384)	(0.741)	(0.374)	(1.037)	(0.656)
Eth Frac^2 (F&L)	-4.392**		-4.288**		-4.908**	
	(0.672)		(0.673)		(1.007)	
ldg	0.436*	-12.58**	0.374*	-12.39**	1.169**	-20.61**
	(0.191)	(1.704)	(0.192)	(1.621)	(0.288)	(3.585)

Continued on next page

Table 6 – *Continued from previous page*

	(1)		(2)		(3)	
	ZIP		ZIP		ZIP	
	Outcome	Inflation	Outcome	Inflation	Outcome	Inflation
phi	-0.094	0.596**	-0.103	0.589**	-0.103	1.246**
	(0.091)	(0.091)	(0.098)	(0.088)	(0.105)	(0.185)
nhi	0.224*	-0.710**	0.195**	-0.774**	0.381**	-0.261
	(0.072)	(0.210)	(0.072)	(0.225)	(0.097)	(0.244)
Controls						
log Population	0.126**	-0.966**	0.104*	-0.943**	0.233**	-1.291**
	(0.049)	(0.086)	(0.048)	(0.081)	(0.097)	(0.198)
Cold War	-0.117		-0.095		-0.225*	
	(0.073)		(0.075)		(0.095)	
Africa	0.126		0.172		0.630**	
	(0.096)		(0.096)		(0.143)	
Constant	-2.768**	5.591**	-2.788**	5.128**	0.389	13.058**
	(0.742)	(1.091)	(0.754)	(1.000)	(1.855)	(2.997)
Observations	4460		4460		2478	
Zero Observations	3543		3543		1910	
Log likelihood	-2335.26		-2341.71		-1308.27	
Vuong test	8.16		8.13		7.13	

Notes: Dependent variable: Conflict prevalence; Standard errors in parentheses; Significance levels:** p<0.01,* p<0.05,† p<0.1; ldg = largest discriminated ethnic group, phi = positive horizontal inequality (relative gap between mean national income and income level of the richest group), nhi = negative horizontal inequality (relative gap between mean national income and income level of the poorest group)

Table 7: Additional Results of ZIP Model: Varying in Income and Democracy

	(1)		(2)		(3)	
	ZIP		ZIP		ZIP	
	Outcome	Inflation	Outcome	Inflation	Outcome	Inflation
Opportunity						
log real GDP					-0.066*	0.403**
					(0.032)	(0.056)
RGDPPC Growth					-2.262**	0.774
					(0.457)	(0.822)
Pri Exports/GDP	-2.212**		-1.838**		-1.605*	
	(0.667)		(0.686)		(0.739)	
Pri Exports/GDP^2	1.958		1.618		1.522	
	(1.129)		(1.147)		(1.253)	
Urbanisation Rate	-0.103*	1.217**				
	(0.049)	(0.242)				
Life Expectancy			-0.887**	1.861**		
			(0.233)	(0.575)		
ΔLife Expectancy			-7.353**	8.988		
			(2.280)	(6.213)		
log % Mountains	0.048*		0.075**		0.079**	
	(0.029)		(0.024)		(0.026)	
Grievance						
Polity Index	0.034**	0.277**	0.032**	0.102**		
	(0.006)	(0.058)	(0.007)	(0.028)		
Polity Index^2	-0.006**		-0.008**			
	(0.001)		(0.001)			
Democracy					-0.420**	1.123**
					(0.099)	(0.186)
Eth Frac (F&L)	5.702**	0.253	6.011**	1.116**	5.859**	1.346**

Continued on next page

Table 7 – *Continued from previous page*

	(1)		(2)		(3)	
	ZIP		ZIP		ZIP	
	Outcome	Inflation	Outcome	Inflation	Outcome	Inflation
	(0.658)	(0.472)	(0.669)	(0.403)	(0.733)	(0.372)
Eth Frac^2 (F&L)	-4.570**		-4.710**		-4.526**	
	(0.593)		(0.603)		(0.674)	
ldg	1.134**	-17.76**	0.767*	-11.63**	0.476**	-12.89**
	(0.139)	(2.687)	(0.187)	(2.006)	(0.181)	(1.645)
phi	-0.199*	1.153**	-0.186*	0.717**	-0.178*	0.575**
	(0.077)	(0.167)	(0.081)	(0.089)	(0.081)	(0.075)
nhi	0.185**	-0.652*	0.207**	-0.439†	0.203**	-0.773**
	(0.064)	(0.268)	(0.074)	(0.283)	(0.072)	(0.242)
Controls						
log Population	0.093**	-0.703**	0.031	-0.642**	0.072	-0.967**
	(0.294)	(0.085)	(0.032)	(0.089)	(0.046)	(0.077)
Cold War	-0.197**		-0.239**		-0.222**	
	(0.064)		(0.072)		(0.068)	
Africa	0.057		-0.066		0.200*	
	(0.083)		(0.088)		(0.095)	
Constant	-3.802**	4.673**	0.574	1.667	-2.647**	5.228**
	(0.549)	(1.350)	(1.048)	(2.327)	(0.701)	(0.865)
Observations	5083		4998		4446	
Zero Observations	4018		3945		3528	
Log likelihood	-2775.87		-2730.70		-2342.33	
Vuong test	9.47		7.97		8.46	

Notes: Standard errors in parentheses; Significance levels:** p<0.01,* p<0.05,† p<0.1; ldg = largest discriminated ethnic group, phi = positive horizontal inequality (relative gap between mean national income and income level of the richest group), nhi = negative horizontal inequality (relative gap between mean national income and income level of the poorest group)

southern africa labour and development research unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy. In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study (2001-) and the Financial Diaries Project.

www.saldru.uct.ac.za
Level 3, School of Economics Building, Middle Campus, University of Cape Town
Private Bag, Rondebosch 7701, Cape Town, South Africa
Tel: +27 (0)21 650 5696
Fax: +27 (0) 21 650 5797
Web: www.saldru.uct.ac.za