Supporting Information Appendix for

"Forecasting Civil Conflict with Zero Inflated Count Models"

Overview

This supplemental appendix first presents a selection of in-sample and out-of-sample ZINB and NB coefficient estimates, standard errors, and model-fit statistics for both the government conflict and rebel conflict models. The supplemental appendix then reports (i) a full table of the classification statistics referenced in the main text, (ii) the 29 countries that are included in the main paper's analysis, (iii) summary statistics and histograms for the main government conflict and rebel conflict dependent variables, and (iv) the formulas used for the classification matrix statistics reported in the main paper.

Estimation Model Results

Table A.1 presents the in-sample (i.e. 1997-2004) coefficient estimates, standard errors, and model fit statistics for the main paper's NB and ZINB models of government conflict and rebel conflict. Following an interpretation of these coefficient estimates and model fit statistics, this section presents a Table of equivalent ZINB and NB models that were run on the entire 1997-2010 sample (Table A.2) in order to demonstrate that the findings reported in Table A.1 are indeed generalizable across the paper's full sample.

Beginning first with the count stages of the government conflict models in the left-hand columns of Table A.1, observe that, for both the ZINB and NB models, higher recent levels of ln government conflicts ($t-1$ to $t-2$) are associated with higher levels of current government conflict at least at the $p < .05$ level, as are higher levels of ln rebel conflict$_{t-1}$. However, ln

rebel conflict$_{t-3}$ is not statistically distinguishable from zero in either model. Moreover, ln government conflict$_{t-3}$ and ln rebel conflict$_{t-2}$ are positive and statistically significant only within the NB model. Overall these findings suggest that the positive reciprocal relationship between past and future levels of government-to-rebel conflict diminishes sharply over time, whereas the inertial attributes of government conflict are relatively more persistent. Moreover, the generally larger NB coefficient estimates may also indicate that—by not accounting for zero inflation—NB models of government conflict overestimate one's coefficient estimates and standard errors, which if further corroborated elsewhere, would be strong support for the contentions made in the main paper. Regarding the three count-stage controls, GDP growth is not statistically significant in either of the government conflict models, while ln GDPpc and ln population are only significant within the NB model. For the NB model, these two latter control variables suggest that countries with (i) lower levels of development or (ii) higher populations are likely to experience more frequent conflict, which is intuitive.

Turning next to the zero inflation stage of the 1997-2004 government conflict models, one can see that—in support of hypothesis 2—all lagged values of ln government conflict are negatively associated (at the $p < .01$ level) with the likelihood that a zero-observation belongs in the hypothesized "zero-only" regime. The same can be said for the coefficient estimates of the lagged values of ln rebel conflict, with the exception of the coefficient estimate for ln rebel conflict$_{t-3}$. Hence, these inflation-stage results suggest that zero-observations that have experienced higher frequencies of recent civil conflict are more likely to be count-stage zeroes rather than observations that could never experience civil conflict. On the other hand, current peace-observations that have experienced little to no recent conflicts are more likely to be structural zeroes, rather than zero-cases that could have reasonably experienced conflict under

different circumstances. Ln GDPpc is positive and significant in the ZINB inflation stage which indicates that higher levels of development decrease the likelihood that a country will ever experience violent government conflicts targeting rebel groups. In sum, the ZINB and NB government-conflict models in Table A.1 suggest that past values of government and rebel initiated material conflict are positively associated with current monthly frequencies of government initiated conflicts, although the causal pathways and estimated relationships therein tend to differ in magnitude and precision.

Table A.1: NB and ZINB Models of Government and Rebel Initiated Conflict, 1997-2004

	NB Gov.	**ZINB Gov.**	**NB Reb.**	**ZINB Reb.**
Ln Gov. Conflict$_{t-1}$	0.355**	0.279**	0.143*	0.165**
	(0.063)	(0.048)	(0.066)	(0.053)
Ln Gov. Conflict$_{t-2}$	0.129*	0.131*	0.026	-0.034
	(0.065)	(0.051)	(0.069)	(0.054)
Ln Gov. Conflict$_{t-3}$	0.225**	0.101	0.168*	0.125*
	(0.063)	(0.052)	(0.066)	(0.053)
Ln Rebel Conflict$_{t-1}$	0.364**	0.200**	0.556**	0.352**
	(0.060)	(0.044)	(0.061)	(0.044)
Ln Rebel Conflict$_{t-2}$	0.270**	0.059	0.342**	0.136**
	(0.062)	(0.048)	(0.063)	(0.050)
Ln Rebel Conflict$_{t-3}$	0.005	-0.032	0.066	0.017
	(0.062)	(0.050)	(0.063)	(0.049)
Ln GDPpc	-0.314**	-0.061	-0.117**	-0.012
	(0.040)	(0.052)	(0.033)	(0.041)
Ln Population	0.247**	0.015	0.170**	-0.017
	(0.023)	(0.026)	(0.021)	(0.024)
GDP Growth	0.004	0.019	-0.066**	-0.044**
	(0.012)	(0.011)	(0.010)	(0.012)
Count Constant	-2.889**	0.734	-2.427**	1.254**
	(0.473)	(0.528)	(0.430)	(0.481)
(Log) Theta	0.733**	0.534**	0.634**	0.431**
	(0.054)	(0.088)	(0.047)	(0.087)
Ln Gov. Conflict$_{t-1}$		-0.636**		-0.200
		(0.202)		(0.200)
Ln Gov. Conflict$_{t-2}$		-0.488*		-1.034**
		(0.212)		(0.303)
Ln Gov. Conflict$_{t-3}$		-0.975**		-0.510*
		(0.244)		(0.226)

Ln Rebel Conflict$_{t-1}$		-0.666**		-0.810**
		(0.170)		(0.169)
Ln Rebel Conflict$_{t-2}$		-0.838**		-0.998**
		(0.202)		(0.223)
Ln Rebel Conflict$_{t-3}$		-0.199		-0.488*
		(0.248)		(0.208)
Ln GDPpc		0.180*		-0.021
		(0.217)		(0.066)
Inflation Constant		-1.211**		-0.281
		(0.605)		(0.517)
Log-likelihood	-2393	-2174	-2618	-2362
AIC	4808.9	4386.4	5257.1	4761.3

Note: N=2,418. ** $p < .01$; * $p < .05$; values in parentheses are standard errors

Similar results are found for the 1997-2004 rebel conflict models reported in Table A.1. For instance, ln government conflict$_{t-1}$ is positive and significantly related to rebel conflict in both of the rebel conflict models. However, in the count stages of the ZINB and NB rebel conflict models in Table A.1, the coefficient estimates for ln government conflict$_{t-2}$ and ln government conflict$_{t-3}$ are not consistently significant. In fact, although generally positive, these two variables are insignificant and occasionally negative-in-sign within these ZINB and NB rebel-conflict models, perhaps suggesting that neither variable has a robust relationship with rebel conflict. Turning to the lagged ln rebel conflict outcome-stage variables, note that ln rebel conflict$_{t-1}$ and ln rebel conflict$_{t-2}$ are positive and significant across both models of rebel conflict, implying that increases in past values of rebel initiated conflict have a positive effect on the frequency of rebel conflict. As above, ln GDPpc and ln population are significant (only) in the NB model, which again suggests that (i) higher levels of development and (ii) smaller populations each decrease the frequency by which countries experience conflict. Additionally, across both the ZINB and NB rebel conflict models, one finds here that GDP growth is negative and significant. This finding implies that higher levels of economic growth lead to lower frequencies of rebel initiated conflicts.

Within the inflation stage of the rebel conflict ZINB model, all lagged values of ln rebel conflict and ln government conflict are negative and significant, save for ln government conflict$_{t-1}$. The former results suggest that increases in past levels of government and rebel initiated civil conflict generally decrease the probability that a peace observation is from the "zero-only" d.g.p., and increase that observation's likelihood of coming from the conflict-count d.g.p. On the other hand ln GDPpc is not significant in the rebel conflict inflation stage, suggesting that development has little effect on preventing a country from ever experiencing a violent rebel-initiated conflict against government actors. Lastly, as above, observe in Table A.1 that the NB model tends to overestimate the count-stage coefficient estimates and standard errors, which is consistent with the expectations of zero inflation, as well as with the results reported for the Vuong tests and AICs above. Hence, for the two rebel-conflict models in Table A.1, higher (lower) past levels of government and rebel initiated conflict are generally associated with higher (lower) current levels of rebel initiated conflict at statistically significant levels, although the predicted relationships for the NB and ZINB models diverge in both the precision of their estimates and the substantive magnitude of their estimated relationships.

Finally, Table A.2 reports coefficient estimates and standard errors for a set of equivalent ZINB and NB models of government conflict and rebel conflict that were run on the entire 1997-2010 sample. As one can see in Table A.2, the coefficient estimates, significance levels, and general findings discussed above for the primary (1997-2004) sample are comparable to those found for (1997-2010) ZINB and NB models of government conflict and rebel conflict. Hence the main paper's (1997-2004) findings and analysis do not appear to rest upon any idiosyncratic features found within this training dataset.

Table A.2: NB and ZINB Models of Government and Rebel Initiated Conflict, 1997-2010

	NB Gov.	**ZINB Gov.**	**NB Reb.**	**ZINB Reb.**
Ln Gov. Conflict$_{t-1}$	0.384**	0.277**	0.104*	0.136**
	(0.046)	(0.035)	(0.047)	(0.039)
Ln Gov. Conflict$_{t-2}$	0.226**	0.169**	0.109*	0.046
	(0.047)	(0.037)	(0.048)	(0.040)
Ln Gov. Conflict$_{t-3}$	0.276**	0.145**	0.077	0.004
	(0.046)	(0.038)	(0.047)	(0.039)
Ln Rebel Conflict$_{t-1}$	0.297**	0.183**	0.521**	0.317**
	(0.045)	(0.035)	(0.046)	(0.034)
Ln Rebel Conflict$_{t-2}$	0.131**	0.032	0.359**	0.186**
	(0.046)	(0.036)	(0.047)	(0.037)
Ln Rebel Conflict$_{t-3}$	0.049	-0.002	0.179**	0.117**
	(0.046)	(0.038)	(0.046)	(0.037)
Ln GDPpc	-0.271**	-0.107*	-0.114**	-0.044
	(0.030)	(0.041)	(0.026)	(0.033)
Ln Population	0.291	0.014	0.222	0.014
	(0.018)	(0.021)	(0.017)	(0.020)
GDP Growth	-0.017	0.011	-0.063**	-0.033**
	(0.009)	(0.009)	(0.008)	(0.010)
Count Constant	-3.905**	0.965*	-3.476**	0.754
	(0.372)	(0.462)	(0.344)	(0.410)
(Log) Theta	0.775**	0.503**	0.716**	0.424**
	(0.044)	(0.069)	(0.042)	(0.068)
Ln Gov. Conflict$_{t-1}$		-0.920**		-0.217
		(0.172)		(0.173)
Ln Gov. Conflict$_{t-2}$		-0.979**		-0.964**
		(0.183)		(0.220)
Ln Gov. Conflict$_{t-3}$		-1.110**		-0.686**
		(0.201)		(0.194)
Ln rebel Conflict$_{t-1}$		-0.527**		-0.830**
		(0.147)		(0.144)
Ln Rebel Conflict$_{t-2}$		-0.548**		-1.071**
		(0.173)		(0.187)
Ln Rebel Conflict$_{t-3}$		-0.367*		-0.628**
		(0.183)		(0.175)
Ln GDPpc		0.055		-0.050
		(0.069)		(0.055)
Inflation Constant		-0.760		-0.233
		(0.503)		(0.442)
Log-likelihood	-4015	-3671	-4259	-3886
AIC	8052.5	7380.2	8540.1	7809.3

Note: N= 4,950. ** $p < .01$; * $p < .05$; values in parentheses are standard errors

Classification Matrix Formulas

$$Sensitivity = \frac{number\ of\ True\ Positives}{number\ of\ True\ Positives + number\ of\ False\ Negatives}$$

$$Specificity = \frac{number\ of\ True\ Negatives}{number\ of\ True\ Negatives + number\ of\ False\ Positives}$$

$$Pos.\ Predicted\ Value = \frac{number\ of\ True\ Positives}{number\ of\ True\ Positives + number\ of\ False\ Positives}$$

$$Neg.\ Predicted\ Value = \frac{number\ of\ True\ Negatives}{number\ of\ True\ Negatives + number\ of\ False\ Negatives}$$

$$Correctly\ Classified = \frac{number\ of\ True\ Positives + number\ of\ True\ Negatives}{number\ of\ cases}$$

Table A.3: Asian and Oceanic Countries Included in the 1997-2010 Sample

Countries	
Australia	Mongolia
Bangladesh	Nepal
Bhutan	New Zealand
Burma	North Korea
Cambodia	Papua New Guinea
China	Philippines
Comoros	Russia
Fiji	Singapore
India	Solomon Islands
Indonesia	South Korea
Japan	Sri Lanka
Laos	Taiwan
Madagascar	Thailand
Malaysia	Vietnam
Mauritius	

Table A.4: Summary Statistics for Monthly Government & Rebel Initiated Conflicts, 1997-2010

	Observations	Mean	Std. Dev.	Variance	Minimum	Maximum
Gov. Initiated Conflicts	5,040	1.89	6.82	46.56	0	98
Rebel Initiated Conflicts	5,040	1.81	6.10	37.19	0	126

Figure A.1: Monthly Frequencies of Rebel & Government Initiated Conflicts, 1997-2010