Supporting Information Appendix for

"Forecasting Civil Conflict with Zero-Inflated Count Models"

Overview

This supplemental appendix first reports and interprets the main (in sample and out-of-sample) ZINB and NB coefficient estimates, standard errors, and model-fit statistics for our *government conflict* and *rebel conflict* models. We then present a tables that report (i) the 29 countries that are included in our analysis and (ii) summary statistics for our main *government conflict* & *rebel conflict* dependent variables. Finally, we report the formulas used for our classification matrix statistics.

Estimation Model Results

Table A.1 below reports the in-sample (i.e. 1997-2004) coefficient estimates, standard errors, and model fit statistics for our main NB and ZINB models of *government conflict* and *rebel conflict*. Following an interpretation of these coefficient estimates and model fit statistics, we conclude this section by briefly reporting a Table of equivalent ZINB & NB models run on our entire 1997-2010 sample (Table A.2) in order to demonstrate that the findings reported in Table A.1 are indeed generalizable across our full sample.

Beginning first with the count stages of the $government\ conflict_{it}$ models in the left-hand columns of Table A.1, we can see that for both the ZINB and NB models, higher recent levels of $ln\ government\ conflicts_{it}$ ($t-1$ to $t-2$) are associated with higher levels of current $government\ conflict_{it}$ at least at the $p < .05$ level, as are higher levels of $ln\ rebel\ conflict_{it-1}$. However, $ln\ rebel\ conflict_{it-3}$ is not statistically distinguishable from zero in either model. Moreover, $ln\ government\ conflicts_{it-3}$ and $ln\ rebel\ conflict_{it-2}$ 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—our NB models of *government* $conflict_{it}$ overestimate our coefficient estimates and standard errors, which if further corroborated below, would be strong support for hypothesis 1. Regarding our three count-stage controls, *GDP growth* is not statistically significant in either of the *government* $conflict_{it}$ models, while *ln* GDP_{pc} and *ln population* are only significant within our 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* $conflict_{it}$ models, we can see that—in support of hypothesis 2—all lagged values of *ln government* $conflict_{it}$ are negatively associated (at the $p < .01$ level) with the likelihood that a zero-observation belongs in our hypothesized "zero-only" regime. The same can be said for the coefficient estimates of the lagged values of *ln rebel* $conflict_{it}$, with the exception of the coefficient estimate for *ln rebel* $conflict_{it-3}$. Hence, our 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* GDP_{pc} is positive and significant in our 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.

[*Insert Table A.1 about here*]

We find similar results for the 1997-2004 *rebel* $conflict_{it}$ models reported in Table A.1. For instance, *ln government* $conflict_{it-1}$ is positive and significantly related to *rebel* $conflict_{it}$ in

both of our *rebel conflict*$_{it}$ models. However, in the count stages of the ZINB and NB *rebel conflict*$_{it}$ models in Table A.1, the coefficient estimates for *ln government conflict*$_{it-2}$ and *ln government conflict*$_{it-3}$ 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*$_{it}$. Turning to the lagged *ln rebel conflict*$_{it}$ outcome-stage variables, we can note that *ln rebel conflict*$_{it-1}$ and *ln rebel conflict*$_{it-2}$ 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*$_{it}$.[1] As above, *ln GDP*$_{pc}$ and *ln population* are significant (only) in our 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 our ZINB and NB *rebel conflict*$_{it}$ models, we find 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*$_{it}$ ZINB model, all lagged values of *ln rebel conflict*$_{it}$ and *ln government conflict*$_{it}$ are negative and significant, save for *ln government conflict*$_{it-1}$. 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 GDP*$_{pc}$ is not significant in our *rebel conflict*$_{it}$ 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, we can note in Table A.1 that the NB model tends to overestimate our count-stage coefficient estimates and standard errors, which is consistent with our 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

[1] Although *ln rebel conflict*$_{it-3}$ is not statistically significant.

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

[*Insert Table A.2 about here*]

Additional Supporting Information

[*Insert Tables A.3 and A.4 about here*]

Classification Matrix Formulas

$$Sensitivity = \frac{number\ of\ True\ Positives}{number\ of\ True\ Positives + number\ of\ False\ Negatives} \tag{1}$$

$$Specificity = \frac{number\ of\ True\ Negatives}{number\ of\ True\ Negatives + number\ of\ False\ Positives} \tag{2}$$

$$Pos.\ Predictive\ Value = \frac{number\ of\ True\ Positives}{number\ of\ True\ Positives + number\ of\ False\ Positives} \tag{3}$$

$$Neg.\ Predictive\ Value = \frac{number\ of\ True\ Negatives}{number\ of\ True\ Negatives + number\ of\ False\ Negatives} \tag{4}$$

$$Correctly\ Classified = \frac{number\ of\ True\ Positives + number\ of\ True\ Negatives}{number\ of\ cases} \tag{5}$$

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

	NB Gov.	**ZINB Gov.**	**NB Reb.**	**ZINB Reb.**
$Ln\ Gov.\ Conflict_{it-1}$	0.355**	0.279**	0.143*	0.165**
	(0.063)	(0.048)	(0.066)	(0.053)
$Ln\ Gov.\ Conflict_{it-2}$	0.129*	0.131*	0.026	-0.034
	(0.065)	(0.051)	(0.069)	(0.054)
$Ln\ Gov.\ Conflict_{it-3}$	0.225**	0.101	0.168*	0.125*
	(0.063)	(0.052)	(0.066)	(0.053)
$Ln\ Reb.\ Conflict_{it-1}$	0.364**	0.200**	0.556**	0.352**
	(0.060)	(0.044)	(0.061)	(0.044)
$Ln\ Reb.\ Conflict_{it-2}$	0.270**	0.059	0.342**	0.136**
	(0.062)	(0.048)	(0.063)	(0.050)
$Ln\ Reb.\ Conflict_{it-3}$	0.005	-0.032	0.066	0.017
	(0.062)	(0.050)	(0.063)	(0.049)
$Ln\ GDP_{pc}$	-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)
$Ln\ Gov.\ Conflict_{it-1}$		-0.636**		-0.200
		(0.202)		(0.200)
$Ln\ Gov.\ Conflict_{it-2}$		-0.488*		-1.034**
		(0.212)		(0.303)
$Ln\ Gov.\ Conflict_{it-3}$		-0.975**		-0.510*
		(0.244)		(0.226)
$Ln\ Reb.\ Conflict_{it-1}$		-0.666**		-0.810**
		(0.170)		(0.169)
$Ln\ Reb.\ Conflict_{it-2}$		-0.838**		-0.998**
		(0.202)		(0.223)
$Ln\ Reb.\ Conflict_{it-3}$		-0.199		-0.488*
		(0.248)		(0.208)
$Ln\ GDP_{pc}$		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. ** indicates $p < .01$; * indicates $p < .05$; values in parentheses are standard errors

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

	NB: Gov. Conflict	**ZINB: Gov. Conflict**	**NB: Reb. Conflict**	**ZINB: Reb. Conflict**
Ln Gov. Conflict$_{it-1}$	0.444**	0.276**	0.129**	0.119**
	(0.049)	(0.034)	(0.050)	(0.038)
Ln Gov. Conflict$_{it-2}$	0.282**	0.168**	0.198**	0.061
	(0.050)	(0.037)	(0.051)	(0.040)
Ln Gov. Conflict$_{it-3}$	0.354**	0.157**	0.094	-0.012
	(0.049)	(0.037)	(0.050)	(0.038)
Ln Reb. Conflict$_{it-1}$	0.299**	0.168**	0.538**	0.348**
	(0.047)	(0.033)	(0.047)	(0.032)
Ln Reb. Conflict$_{it-2}$	0.154**	0.046	0.335**	0.171**
	(0.049)	(0.034)	(0.048)	(0.037)
Ln Reb. Conflict$_{it-3}$	0.041	-0.016	0.219**	0.125**
	(0.049)	(0.036)	(0.048)	(0.036)
Count Constant	-0.821**	0.562**	-0.735**	0.520**
	(0.034)	(0.049)	(0.033)	(0.050)
Theta	0.598**		0.594**	
	(0.033)		(0.033)	
Log Theta		0.473**		0.315**
		(0.068)		(0.066)
Ln Gov. Conflict$_{it-1}$		-0.883**		-0.251
		(0.145)		(0.156)
Ln Gov. Conflict$_{it-2}$		-0.909**		-0.985**
		(0.162)		(0.219)
Ln Gov. Conflict$_{it-3}$		-0.957**		-0.807**
		(0.170)		(0.199)
Ln Reb. Conflict$_{it-1}$		-0.527**		-0.804**
		(0.131)		(0.133)
Ln Reb. Conflict$_{it-2}$		-0.509**		-0.980**
		(0.149)		(0.171)
Ln Reb. Conflict$_{it-3}$		-0.463**		-0.653**
		(0.160)		(0.162)
Inflation Constant		-0.232		-0.718**
		(0.188)		(0.260)
LogLikelihood	-4,598	-4,050	-4,780	-4,290
AIC	9,213	8,139	9,467	8,616
Number of Obs.	4,950	4,950	4,950	4,950

Note: ** indicates $p < .01$; * indicates $p < .05$; values in parentheses are standard errors

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 and Rebel Initiated Domestic Material Conflicts, 1997-2010

	Observations	Mean	Std. Dev.	Variance	Minimum	Maximum
Government 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