Online Appendix for "The Prospects of Punishment and the Strategic Escalation of Civil Conflicts" (Not for publication)

Casey Crisman-Cox*

June 23, 2020

Contents

*Texas A&M University. email: c.crisman-cox@tamu.edu.

A Summary statistics and validity checks

In this appendix, I report summaries of the main independent variables. Table A.1 reports the summary statistics and sources for all the independent variables used in the main model.

Table A.1: Independent Variable Summaries

Variable	Min	Mean	Max	Source
New state	0.00	0.05	1.00	COW
Population (logged)	2.94	9.91	14.06	NMC
Ethnic Frac.	0.00	0.50	1.00	Fearon and Laitin (2003)
Territorial	−10.00	−0.61	10.00	UCDP
Polity2	0.00	0.47	0.93	Polity IV
Oil	0.00	0.14	1.00	World Bank & Fearon and Laitin (2003)
Mount. Terrain (logged)	0.00	2.56	4.56	Fearon and Laitin (2003)
GDP pc (logged)	−1.61	0.99	4.68	Penn World Table
Mil. Per. pc (logged)	0.00	4.52	8.52	NMC
Other groups (logged)	0.00	2.04	3.71	CONIAS & MARS

As a check of the data, I also consider a Heckman-selection model to see if we can reproduce some of the findings from Fearon and Laitin (2003). This analysis serves as a check on the face-validity of the data. For the selection model we use a non-violent dispute (outcomes BD and CD) as the first step's dependent variable and a violent conflict as the second step's outcome (as in Bartusevičius and Gleditsch 2018). All of the variables from the main model are included with population and new state both only entering the selection equation to satisfy the exclusion restriction. Note that this is not a replication exercise as I use different variables and a different model specification, but rather this is just to see if this data provides results that match these canonical results on civil conflict onset. The results of this exercise are presented in Table A.2.

In the selection model we see a few results that line up with the standard Fearon and Laitin (2003) results. Notably, we see that being an oil exporter and having a large proportion of mountainous terrain is associated with a larger risk of conflict, while proxies for state power (GDP per capita and military personnel per capita) are associated with a lower risk of conflict. Overall, these basic results match with standard civil conflict results and provide some validity to the CONIAS data. For more of these types of validation exercises see Bartusevičius and Gleditsch (2018).

Table A.2: Validating the data with previous results

	Dependent variable:	
	Non-violent dispute	Violent conflict
New state	1.01***	
	(0.13)	
Population	0.02	
	(0.05)	
Oil	0.15	0.14*
	(0.10)	(0.08)
GDP per capita	−0.05	−0.09***
	(0.05)	(0.03)
Democracy	0.01**	−0.002
	(0.01)	(0.004)
Democracy squared	−0.0003	0.001
	(0.001)	(0.001)
% Mountainous terrain	−0.004	0.05**
	(0.03)	(0.02)
Military personnel per capita	−0.02	−0.03*
	(0.04)	(0.02)
Ethnic fractionalization	−0.14	0.18*
	(0.14)	(0.10)
Other groups	0.03	−0.001
	(0.05)	(0.04)
Constant	−1.64***	0.84***
	(0.33)	(0.19)
Observations	3,513	
Log Likelihood	−878.60	
ρ	−0.20	
Note:	*p<0.1; **p<0.05; ***p<0.01	

B Episodic data

In this appendix, I consider a different approach to aggregating the data. Specifically, here I look at the episodes in terms of a government-group-episode rather than the government-group-decade. Each government-group pair now only enters the data once and only the most extreme outcome of the interaction is considered. The main effect here is to reduce or limit the number of status quo observations.

The results from the episodic data are reported in Table B.1. Here we see that the estimates are largely unchanged, although there are some minor differences. Most importantly however,

Table B.1: Punishment estimates for episodic data

	Estimate	Standard Error
Constant	-2.08	0.69
Democracy	0.04	0.04
Democracy squared	-0.01	0.01
Other groups	-1.15	0.50
Military personnel per captia (log)	-0.02	0.12
Separatist dispute	-0.05	0.39
Log-likelihood = -784.46		
Observations = 406		

the punishment parameters in $\bar{a}$ are roughly the same magnitude and direction. As with the dyad-decade data, we see that on average groups expect to be punished for backing down in a dispute and that these punishments are relatively sizable compared to the group's other payoff parameters. Interestingly the coefficient on the number of other groups is notably larger and is significant at conventional levels. However, far fewer observations in this model satisfy the sensible payoff restrictions, which suggests that the decade level aggregation provides better model fit.

C Fixed-effects approaches

In this appendix, I consider both country-specific and then group-specific fixed-effects approaches to specifying the punishment parameters $\bar{a}$. The pseudo-likelihood routines exhibited signs of numerical instability, which made standard error estimation unreliable. As such, no standard errors are reported; the values in this section are best thought of as a model calibration exercise.

The country-specific punishments are presented in Figure C.1, while the group-specific results are presented in Figure C.2. A few interesting results are apparent. First, the estimates are overwhelmingly negative, further supporting the findings in the main text. For almost all groups backing down is, on average, a very costly proposition. Second, there is noticeable heterogeneity in the estimates. Some groups expect much harsher penalties than others. This results represents a very interesting venue for future work: where do these differences comes from and what attributes of groups makes larger punishments more likely? Sadly data limitations on group-level covariates preclude such an analysis at this time, although the number of other potential groups still correlates strongly with these measures ($p < 0.05$).

Figure C.1: Calibrated estimates of $\bar{a}$ with country fixed effects

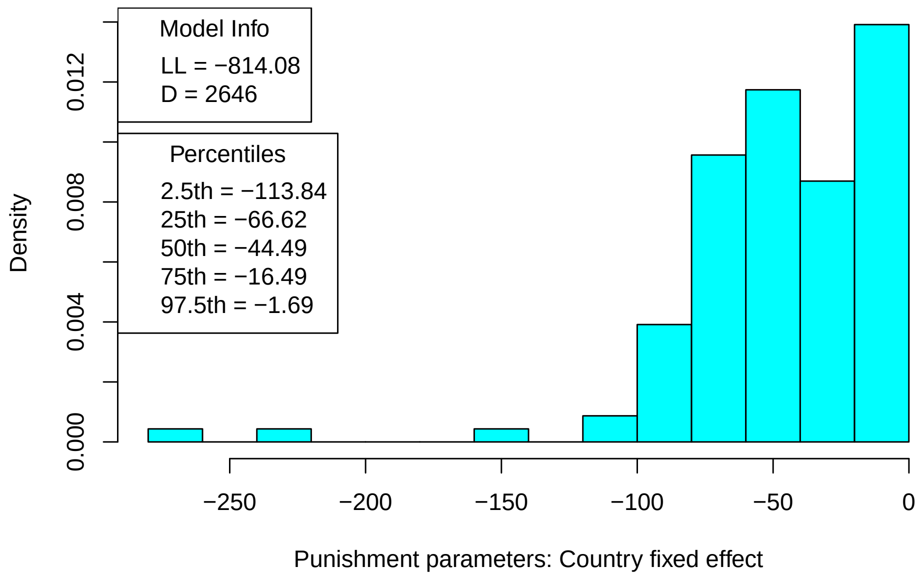
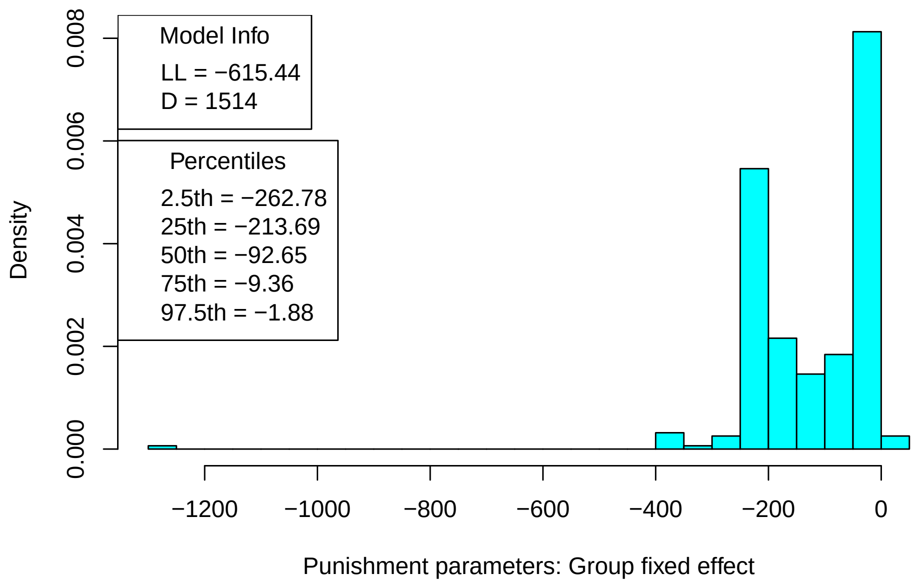


Figure C.2: Calibrated estimates of $\bar{a}$ with group fixed effects



D Additional robustness checks

In this appendix, I consider extensions to the main model with respect to the punishment parameters $\bar{a}$, the sample used, and measurements of key concepts. Across these models we are interested in whether the constant term remains negative and significant and whether the number of other groups remains a significant correlate of the expected punishment.

Table D.1: Additional robustness checks

	Main model	External punishment	Goals	Post-Cold War	Using rugged terrain	No empires	Ambiguous cases
Constant	-2.83**	-2.89**	-2.86**	-2.54**	-2.83**	-3.01**	-2.88**
	(0.33)	(0.37)	(0.41)	(0.76)	(0.41)	(0.47)	(0.38)
Democracy	0.02	0.03	0.03	0.01	0.03	0.04	0.01
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.02)
Democracy squared	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Other groups	-0.14†	-0.19*	-0.18*	-0.56†	-0.20*	-0.20*	-0.14
	(0.08)	(0.08)	(0.08)	(0.33)	(0.08)	(0.09)	(0.10)
Military personnel per captia		0.09	0.09	0.30	0.10	0.05	0.10
		(0.07)	(0.08)	(0.20)	(0.08)	(0.09)	(0.07)
Separatist dispute			-0.06	-0.12	-0.04	-0.02	0.14
			(0.34)	(0.58)	(0.34)	(0.39)	(0.23)
Log L	-864.51	-863.85	-863.57	-298.90	-861.36	-817.22	-1062.38
N Games	2794	2794	2794	940	2765	2741	2865

Notes: $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$
Standard errors in parenthesis

E Estimation details

Recall that the equilibrium of the model can be described as a system of three equations, where p^* is an equilibrium only if $p^* = \Psi(p^*; \theta)$. We now write out the full form of this system of equations:

$$p_c^* = 1 - \Phi\left(\frac{S_R - (1-p_r^*)V_R}{p_r^*} - \bar{W}_R\right)\Phi\left(\frac{S_R - (1-p_r^*)V_R}{p_r^*} - \bar{a}\right) \equiv f(p_r^*; \theta), \tag{E.1}$$

$$p_e^* = \Phi_2\left(\frac{\bar{W}_R - \bar{a}}{\sqrt{2}}, \bar{W}_R - \frac{S_R - (1-p_r^*)V_R}{p_r^*}; \frac{1}{\sqrt{2}}\right) \Big/ f(p_r^*; \theta) \equiv h(p_r^*; \theta), \tag{E.2}$$

and

$$p_r^* = \Phi\left(\frac{h(p_r^*; \theta)\bar{W}_G + (1 - h(p_r^*; \theta))V_G - C_G}{h(p_r^*; \theta)}\right) \equiv g \circ h(p_r^*; \theta). \tag{E.3}$$

In the above, $\Phi(x)$ is the standard normal cumulative distribution function (CDF) and $\Phi_2(x, y; \rho)$ is the standard bivariate normal CDF ($\sigma_x^2 = \sigma_y^2 = 1$) with correlation ρ. Notice that p_r^* completely pins down the equilibrium: Equations E.1 and E.2 are R's best responses to p_r^*, while Equation E.3 is G's best response to $h(p_r^*; \theta)$, as such we can rewrite the system as $p^* = \Psi(p_r^*; \theta)$. Because this is not an original model, I do not derive these expressions of the choice probabilities except to note the probabilities are all formed based on ordinary comparison of the expected utility of an action compared to private information as in nearly all random utility models (e.g., Signorino 1999). More thorough discussions of the derivations can be found in Lewis and Schultz (2003) or Jo (2011).

Let $y_d \in \{SQ, CL, BD, CC\}$ denote the outcome of a group and government dyadic observation $d = 1, \dots, D$. To estimate the parameter vector β, we start by constructing the multinomial log-likelihood

$$\begin{aligned} L(\beta|y) = \sum_{d=1}^{D} & \mathbb{I}(y_d = SQ)\log\left[1 - f(p_{r,d}; \beta)\right] + \mathbb{I}(y_d = CL)\log\left[f(p_{r,d}; \beta)(1 - g(h(p_{r,d}; \beta)))\right] \\ & + \mathbb{I}(y_d = BD)\log\left[f(p_{r,d}; \beta)g(h(p_{r,d}; \beta))(1 - h(p_{r,d}; \beta))\right] \\ & + \mathbb{I}(y_d = CC)\log\left[f(p_{r,d}; \beta)g(h(p_{r,d}; \beta))h(p_{r,d}; \beta)\right], \end{aligned} \tag{E.4}$$

where $p_{r,d}$ is found by solving Eq. E.3 for each observation at every guess of β and $\mathbb{I}(\cdot)$ is the indicator function. Ideally, we want to find the values of β that maximize this log-likelihood. However, it is well known that this game admits multiple equilibria under many reasonable sets of parameters (Crisman-Cox and Gibilisco 2019; Jo 2011), which makes this task less than straightforward.

In particular, Crisman-Cox and Gibilisco (2019) demonstrate that Eq. E.4 is not well defined because multiple equilibria are possible and using Eq. E.4 is problematic for estimation *even if there is a unique equilbrium at the true parameters values.* Specifically, this issue means that any given guess of the parameter vector could be consistent with multiple values of p_r for each observation and can consequently produce multiple log-likelihood values (which is to say that the above log-likelihood is a correspondence rather than a function). As a result of this multiplicity, standard optimization techniques will frequently converge at incorrect estimates (incorrect in the sense of not actually maximizing Eq. E.4). They also note that standard refinements, such as the Intuitive Criterion or regularity, do not provide any relief here as all equilibria of this game will survive the refinements. Likewise, they show that that while other ad-hoc refinements can be employed to eliminate the multiplicity, they will introduce severe discontinuities into the above log-likelihood and make direct optimization of Eq. E.4 substantially less feasible (and the number of discontinuities tends to increase with the number of observations).

As a result of these concerns I use the nested pseudo-likelihood (NPL) estimator they propose to find the parameters of interest.[1] This estimation procedure sidesteps the problems associated with multiple equilibria by assuming that equilibrium selection is a function of the observed covariates. To put this another way, this approach imposes an equilibrium selection rule that is empirical, rather than theoretical, and allows the data to tell us which equilibrium the actors reach. The estimation routine proceeds as follows:

1. The equilibrium choice probabilities $p_{r,d}$ and $p_{e,d}$ are estimated using a flexible method (in this case a random forest) that relates the decision nodes to all the observed covariates. These initial estimates of equilibrium behavior need not, and likely will not, satisfy the equilibrium conditions in Equations 1-3, but this is not an issue as these probabilities only serve as an initial guess in an iterative process.

2. Estimate β by maximizing the log-pseudo-likelihood with the equilibrium quantities fixed to the current estimates of $p_{r,d}$ and $p_{e,d}$.

[1] Specifically, I use their `R` package (`sigInt`) for estimation and analysis.

3. Using the estimates of β from step 2, update the estimates of $p_{r,d}$ and $p_{e,d}$ using best-response functions g and h, respectively.

4. Iterate steps 2 and 3 until convergence.

The log-pseudo-likelihood from step 2 is given by:

$$\begin{aligned} PL(\beta|y,\hat{p}) = \sum_{d=1}^{D} & \mathbb{I}(y_d = SQ) \log\left[1 - f(\hat{p}_{r,d};\beta)\right] + \mathbb{I}(y_d = CL) \log\left[f(\hat{p}_{r,d};\beta)(1 - g(\hat{p}_{e,d};\beta))\right] \\ & + \mathbb{I}(y_d = BD) \log\left[f(\hat{p}_{r,d};\beta) g(\hat{p}_{e,d};\beta)(1 - h(\hat{p}_{r,d};\beta))\right] \\ & + \mathbb{I}(y_d = CC) \log\left[f(\hat{p}_{r,d};\beta) g(\hat{p}_{e,d};\beta) h(\hat{p}_{r,d};\beta)\right], \end{aligned} \tag{E.5}$$

where $\hat{p} = (\hat{p}_{e,d}, \hat{p}_{r,d})_{d=1}^{D}$ refers to the current estimates of the choice probabilities from steps 1 and 3. Notice that the equilibrium quantity $p_{r,d}$ is no longer endogenously defined and so it does not have to computed at every optimization step; this means that the indeterminacies/discontinuities in Eq. E.4 are not present in the pseudo-likelihood function. The intuition behind this approach is that if we know the true equilibrium choice probabilities we could fix $p_{r,d}$ to these values in Eq. E.4, which would turn it into a well-behaved and continuous function. However, since we do not know these choice probabilities we estimate them from the observables to generate a feasible estimator. The iterative process means that the NPL estimates will be in equilibrium at convergence and overall the estimation routine will be both better and faster than trying to directly optimize Eq. E.4. For more details see Crisman-Cox and Gibilisco (2019).

References

Bartusevičius, Henrikas and Kristian Skrede Gleditsch. 2018. "A Two-Stage Approach to Civil Conflict: Contested Incompatibilities and Armed Conflict." *International Organization* 73:225–248.

Crisman-Cox, Casey and Michael Gibilisco. 2019. "Estimating Crisis Signaling Games in International Relations: Problems and Solutions." Forthcoming at *Political Science Research and Methods*.

Fearon, James D. and David D. Laitin. 2003. “Ethnicity, insurgency, and civil war.” *American Political Science Review* 97(1):75–90.

Jo, Jinhee. 2011. “Nonuniqueness of the Equilibrium in Lewis and Schultz’s Model.” *Political Analysis* 19(3):351–362.

Lewis, Jeffrey B. and Kenneth A. Schultz. 2003. “Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information.” *Political Analysis* 11(4):345–367.

Signorino, Curtis S. 1999. “Strategic Interaction and the Statistical Analysis of International Conflict.” *American Political Science Review* 93(2):279–297.