Online Appendix for "Democracy, Reputation for Resolve, and Civil Conflict" (Not for publication)

February 10, 2021

Contents

A Estimation and computational details

A.1 Likelihood function

Our goal here is to build a likelihood function to estimate the parameter vector $\theta = (\beta_R, \beta_G, \gamma_R, \gamma_G)$. Three values define the likelihood for each observation. First, there is the question of which side is the possible first quitter (player 1). For a given step in the numerical optimization, this is a function of the current parameter vector θ and the observables (X, Z). Specifically, I define an indicator for when G is the possible quitter:

$$\delta_n^1(\theta|X_n, Z_n) := \begin{cases} 1 & \frac{\log(z_R(Z_n,\gamma_R))}{\lambda_R(X_n,\beta_R)} > \frac{\log(z_G(Z_n,\gamma_G))}{\lambda_G(X_n,\beta_G)} \\ 0 & \text{otherwise.} \end{cases}$$

As the optimization routine searches for a solution, this indicator will change values.

Second, I need to know the conflict's outcome. Specific outcomes, particularly in civil conflict, can be slightly ambiguous. While I discuss the exact operationalization in the data section (main text), it is worth mentioning that terminations here refer to one side offering policy concessions or otherwise backing down from the conflict. Whenever one side backs down I only know that one player exited at time t, while the other player's duration is censored. If the conflict is ongoing, then both durations are censored. I define a variable denoting the three outcomes:

$$\delta^w := \begin{cases} 0 & R \text{ backs down} \\ 1 & G \text{ backs down} \\ 2 & \text{Conflict is ongoing or one side was military defeated} \end{cases}$$

Note that while δ^w is coded prior to estimation, δ^1 is determined from the data and the current parameter values. The final thing I need to know is if the "possible first quitter" actually quits at time 0.[1]

Recall that $n = 1, \ldots, N$ denotes a specific conflict, where one government and one rebel group play the continuous-time bargaining game. Within the conflict we can think of each side first being assigned their type (unresolved or resolved) and then we are interested in their strategies given their type. This creates a log-likelihood where for civil conflict n, the rebel group's contribution to the

[1] The operationalization of duration (particularly length 0 durations) is discussed in the data section and below in Appendix D.4.

likelihood is

$$L_R(\theta|y_n, Z_n, X_n) = \begin{cases} 1 - F_2(y_n|\theta, Z_n, X_n) & \delta_n^1(\theta|X_n, Z_n) = 1, \delta_n^w \geq 1 \\ 1 - F_1(y_n|\theta, Z_n, X_n) & \delta_n^1(\theta|X_n, Z_n) = 0, \delta_n^w \geq 1 \\ f_2(y_n|\theta, Z_n, X_n) & \delta_n^1(\theta|X_n, Z_n) = 1, \delta_n^w = 0 \\ f_1(y_n|\theta, Z_n, X_n) & \delta_n^1(\theta|X_n, Z_n) = 0, \delta_n^w = 0, y_n > 0 \\ F_1(y_n|\theta, Z_n, X_n) & \delta_n^1(\theta|X_n, Z_n) = 0, \delta_n^w = 0, y_n = 0. \end{cases} \tag{A.1}$$

To understand the derivation, recall that the probability of a rebel group quitting before time y_n given that they are unresolved is $\frac{F_R(y_n|\theta, Z_n X_n)}{1 - z_R(Z_n, \gamma_R)}$. The first two lines in Equation A.1 are cases where the rebel duration is right censored. In these cases the likelihood is

$$(1 - z_R(Z_n, \gamma_R)) \left(1 - \frac{F_R(y_n|\theta, Z_n X_n)}{1 - z_R(Z_n, \gamma_R)}\right) + z_R(Z_n, \gamma_R)(1),$$

which simplifies into lines one and two of Equation A.1. Alternatively, recall that F_R reflects the total probability that a rebel group exits at or before y_n. Note that this derivation includes the possibility that resolved types exist in the world and that the first two lines of Equation A.1 are the probability that R is either resolved or is unresolved but has not backed down by y_n.

The same logic defines lines three and four, but these are conflicts where a rebel group is unresolved and drops out at time y. In these cases, the likelihood simplifies to the density associated with this joint event, where f_i is the derivative of F_i with respect to y. Finally, the last line indicates conflicts where rebels are the first quitters and actually drop out at $y = 0$.[2]

The government's contribution to conflict n's likelihood is defined similarly, making the combined log-likelihood

$$L(\theta|y, Z, X) = \sum_{n=1}^{N} \log\left[L_R(\theta|y_n, Z_n, X_n)\right] + \log\left[L_G(\theta|y_n, Z_n, X_n)\right]. \tag{A.2}$$

Note that there is a discontinuity in the log-likelihood introduced by the inequality that defines $\delta_n^1(\theta|X_n, Z_n)$. As such a non-parametric bootstrap is used to generate confidence intervals on estimated parameters.[3]

A.2 Computational details

This appendix discusses the computational details for obtaining point estimates and confidence intervals. The discontinuous log-likelihood explicitly precludes derivative-based numerical optimizers. Monte Carlo analysis (below) suggests that the standard derivative free solver (Nelder-Mead) is

[2] In situations where one side is known to exit the conflict at $y_n = 0$, δ_n^1 is hard coded to reflect the data and avoid a zero-likelihood problem.

[3] Chernozhukov and Hong (2004) suggest that maximum likelihood with a bootstrap is valid when estimating auction models (which this model is a type of) with a discontinuous likelihood.

sufficient for the problem. However to better ensure that the actual optimum is reached, I employ a global optimizer that uses a differential evolution (DE) algorithm to find starting values for the Nelder-Mead solver. The DE algorithm does not require that the objective function be either continuous or differentiable (Mullen, Ardia, Gil, Windover and Cline 2011).

The main concern with the DE optimizer is that there is no predefined way to assess convergence. I run the algorithm in sets of 500,000 iterations, warm restarting as needed, until 100,000 iterations in a row show no noticeable improvement in the objective function. To be precise, there must be 100,000 iterations in a row where the likelihood function is not increased by a factor of at least $\varepsilon \times (|L(\hat{\beta}, \hat{\gamma}; y, X, Z)| + \varepsilon)$, where $\varepsilon \approx 1.5 \times 10^{-8}$ is the square root of the machine tolerance. The reduction factor is the software's default, and I choose 100,000 iterations to provide confidence that these will be good starting values. Once this procedure finishes, I pass the resulting estimates through a Nelder-Mead optimizer to search for additional improvement. Nelder-Mead is also used to bootstrap the standard errors as the global optimization is too computationally expensive to use in a bootstrap.

B Monte Carlo experiments

In this appendix, I detail a Monte Carlo experiment designed to demonstrate the strategic statistical duration model performs as expected. The parameters in the experiment are fixed to the following:

$$\begin{aligned}
\lambda_R &= \exp(-3.0 + 0.5X) \\
\lambda_G &= \exp(-2.7 - 0.5X) \\
z_R &= \text{logit}^{-1}(1 + 0.5X) \\
z_G &= \text{logit}^{-1}(1.5 - 0.5X),
\end{aligned}$$

In all experiments X is drawn from the standard uniform distribution. The vector of true parameters is

$$\theta = (-3.0, 0.5, -2.7, -0.5, 1, 0.5, 1.5, -0.5).$$

These parameters are used to create duration data, y. I follow the log-likelihood in generating the data, such that I generate the rebel durations by drawing

$$y_{R,n} \sim \begin{cases} F_2 & \lambda_{R,n}^{-1} \log(z_{R,n}) \geq \lambda_{G,n}^{-1} \log(z_{G,n}) \\ F_1 & \lambda_{R,n}^{-1} \log(z_{R,n}) < \lambda_{G,n}^{-1} \log(z_{G,}), \end{cases}$$

and likewise for the government.

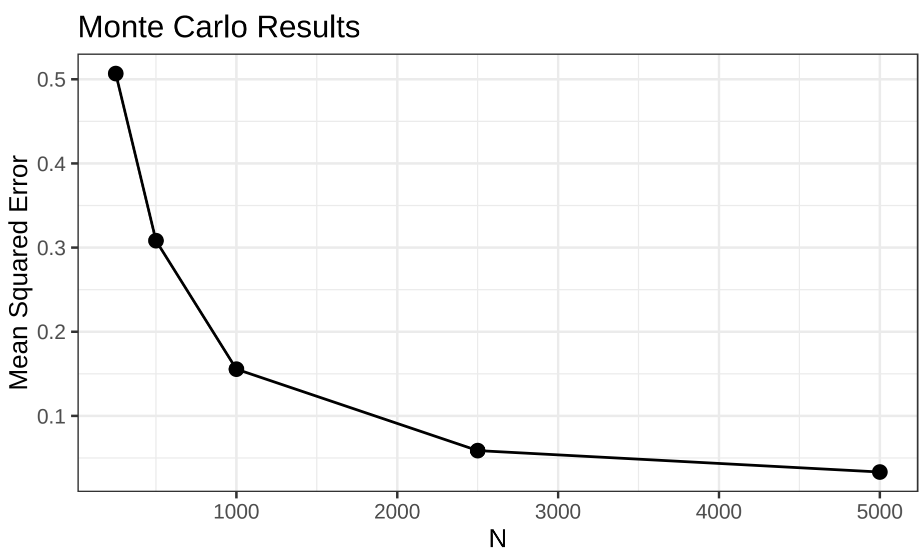
Because the game ends when one player quits, y is a vector defined as $y_n = \min\{y_{R,n}, y_{G,n}\}$. To make sampling easier, I allow F_1 to produce negative durations, which are set to be 0 (as in simulations from Ramsay and Signorino 2009). To include resolved types and other censored observations I find $\bar{z} = \frac{1}{N} \sum_{n=1}^{N} z_{R,n} z_{G,n}$ and then set the proportion of censored conflicts to $\bar{z}$.

This approach provides a reasonable and straightforward way to include the possibility of resolved types and other censoring in the simulation. The simulation results are unchanged if the max of $z_R z_G$ is used in place of the mean.

These parameters thus give us a wide spread over the possible outcomes and allow us to see how the estimator preforms with all of these factors at play. In particular, when the starting values are set to 0 for all parameters the model predicts that in every case the rebel is the i player, meaning that optimizer has to find parameters that flip the value of the inequality. In short, the parameters values provide a good test of the estimator's ability to handle the discontinuities and other difficulties.

The Monte Carlo experiment is conducted at five different sample sizes and each experiment is repeated 250 times. Figure B.1 summarizes the results of the Monte Carlo experiments. Here we see the mean-squared error (MSE) averaged over 8 parameters. Overall, the MSE is decreasing as N increases, which provides confidence in its performance.

Figure B.1: Monte Carlo experiments for the war-of-attrition model



C Summary statistics and conflict episodes

Table C.1 reports the summary statistics and sources for all variables used in the main model.

Table C.1: Independent Variable Summaries

Variable	1st Quartile	Average	3rd Quartile	Source/Measurement
Duration (Months)	4.00	14.00	37.00	UCDP
GDP per capita (logged)	−0.01	0.61	1.43	Penn World Table
Polity 2	−7.00	0.00	7.00	Polity IV
Population (logged)	8.69	9.67	10.80	COW-NMC
Mil. Per. per capita (logged)	3.00	4.17	5.40	COW-NMC
Religious Frac.	0.27	0.44	0.61	Fearon and Laitin (2003)
Ethnic Frac.	0.36	0.70	0.83	Fearon and Laitin (2003)
Mountainous terrain (logged)	1.86	2.44	3.64	Fearon and Laitin (2003)
Coup	0.00	0.13	0.00	Cunningham (2006)
Oil	0.00	0.18	0.00	Fearon and Laitin (2003)
Territorial	0.00	0.41	1.00	UCDP

Note: Reported averages are medians for continuous variables and means for binary variables

There are several advantages to using conflict episodes as the unit of analysis. One advantage is that they capture are relatively short conflicts. As a result, within some countries they will capture demise and rebirth of conflicts, although these cycles do not routinely appear in the data. Additionally, episodes keep more data and require fewer questionable coding choices than aggregating multiple conflict episodes. An additional advantage of using conflict episodes is that they sidestep concerns about how the independent variables or actors themselves change over the course of an episode. For example, there are no leadership changes within more than 80% of the conflict episodes, mitigating concerns about treating the government as unitary and suggesting that leaders who might be concerned about being removed from office are focused on their post-conflict-episode fates. Further, the outcome data can vary across episodes within a country and treating episodes individually reduces the need to aggregate across these and further limit the sample.

The main concern with using UCDP conflict episodes is that the low-inclusion threshold (25 battle deaths in a given year) might lead a lot of low capacity groups to exit and reenter the data prematurely. However, this concern does not appear to be pervasive in the data as most countries experience only 1-3 episodes total. As such, while some groups may exit and re-enter the data, this cycling is not routine in the data. A second concern is that episodes are assumed to be independent to make the estimation problem tractable. This independence assumption is a standard feasibility constraint in structural work (e.g., Crisman-Cox and Gibilisco 2018; Whang, McLean and Kuberski 2013). Future work however should consider better ways to incorporate within country changes that occur over the course of the conflict and across episodes to relax some of the independence assumptions. This more expansive model will also better capture the effects of past conflict outcomes on future reputation.

D Additional results and robustness checks

In this appendix, I present the results from some additional models. The first includes the β estimates that make up the λs for Model 1 (Table 2 of the main text). These estimates are presented in Table D.1, below. The estimated βs are largely similar across the other models and as such are omitted.

Table D.1: Estimates of λ_R and λ_G from Model 1

	Rebel	Government
Intercept	-3.70	-1.92
	$(-5.73, -2.82)$	$(-2.91, 0.18)$
GDP per capita (log)	0.06	-0.43
	$(-0.18, 0.30)$	$(-0.66, -0.09)$
Democracy	0.00	0.05
	$(-0.04, 0.03)$	$(0.01, 0.11)$
Population (log)	-0.15	-0.33
	$(-0.26, 0.06)$	$(-0.53, -0.20)$
Miltary personnel per capita (log)	0.06	0.11
	$(-0.10, 0.29)$	$(-0.20, 0.24)$
Religious Frac.	0.64	-0.13
	$(-0.36, 1.73)$	$(-1.22, 1.53)$
Ethnic Frac.	0.41	1.01
	$(-0.41, 1.37)$	$(0.02, 2.03)$
Mt. Terrain	-0.00	-0.33
	$(-0.16, 0.15)$	$(-0.58, -0.11)$
Coup	0.31	0.13
	$(-0.52, 1.20)$	$(-3.20, 1.13)$
Oil	0.37	0.22
	$(-0.27, 0.86)$	$(-0.61, 1.02)$
Territorial Conflict	-0.05	0.27
	$(-0.38, 0.40)$	$(-0.32, 0.73)$
Log L	-1551.56	
N	397	

Notes: Bootstrapped 95% confidence interval in parentheses

Recall that λ_i is the hazard rate for player i's exit time. As such, positive coefficients mean that higher values of the variable are associated with i exiting the conflict faster, while negative coefficients imply that i persists longer in conflict. There are three points of interest here. First, is the effect that democracy has on unresolved governments: unresolved democracies exit sooner, on average, than unresolved autocracies. This matches the reputation results from the main text, in that unresolved democratic governments realize that it is harder for them to build a reputation for resolve and so they do not invest as much into doing so. Second, GDP per capita is associated with government endurance in the expected direction. This provides some face validity to the results as it suggests wealthier states are willing/able to fight for longer periods of time. Finally, the basic civil war predictors do a much better job explaining government actions than rebel actions. This is perhaps not surprising given the focus of past work on state-level correlates of civil conflict duration and outcomes.

D.1 Additional results from Models 1 & 3

In this appendix, I further explore the models in Table I (main text), starting with Model 1. Here, reputation for resolve is modeled as a constant for all governments and all rebel groups. Overall, both sides think that their opponent is very likely to be resolved.[4] This fits with the empirical record of civil wars and provides some important face validity to the results. Specifically, within the data, 48% of conflicts end with rebels exiting the conflict without changes to the status quo, 23.5% end with the government accepting some change to the status quo, and the remaining 28.5% are either a direct military victory or ongoing (i.e., each side's exit time is censored). As such, in 76.5% of all observations the government never reveals its type, while in about 52% of observations, the rebel group does not reveal its type (these values roughly match the conflict summaries reported by Kreutz (2010)). These high initial reputations for revolve reflect the realities of civil conflict where actors rarely reveal their type.[5]

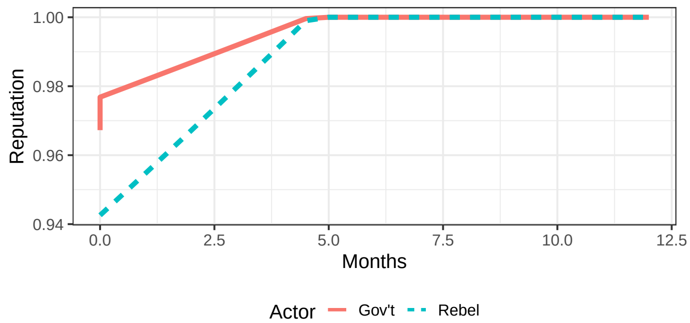
From the point estimates we see that both sides begin conflict with a strong reputation for resolve, but there are a few differences in how they develop. In Figure D.1, we again see that governments begin with a reputation advantage. While the government seems to enjoy a slight advantage in its initial reputation, updating (reputation building) does not occur at the same rate for both actors. Specifically, we see that in an average conflict (all the covariates in λ are set to a mean or median value), rebels tend to build reputation faster than governments. The exception to this is in the first instant of conflict when both sides decide to continue fighting beyond the initial incompatibility. Here, an average governments receives an immediate reputation boost when conflict continues past the initial incompatibility. This suggests that there is some uncertainty about whether governments will decide to fight (i.e., governments are, on average, the possible first quitter) and the rebel group updates strongly when they do. One way to interpret this boost in the state's reputation for resolve is in terms of credibility theory and the prospect of punishment — once a state finds itself in conflict, backing down becomes a costly act. Risking this cost means that the state is much more likely to be the resolved type that does not care about the costs of conflict,

[4] The constants in Model 2 are not directly interpretable in this way, as they refer to governments and rebels in states where the other covariates are all equal to zero.

[5] The government's advantage in initial reputation for resolve appears in approximately 92% of of the bootstrap iterations. While there is little existing theory to suggest which side should have a stronger reputation for resolve heading into conflict, the government advantage here is not unintuitive; governments are, in general, longer lived actors who have more time and resources to put into building a reputation for resolve prior to the start of the conflict process. Scholars overwhelmingly acknowledge that leaders care about and invest in their reputations (e.g, Mercer 1996), which might be reflected in this result.

which results in the initial increase once fighting begins.

Figure D.1: Reputation over time for an average conflict



In addition to understanding how reputations develop over the course of the conflict, we also want to know if a reputation for resolve actually affects the outcomes of conflict. Specifically, we will look at the probability that an unresolved government gives into rebels. Here, we simulate each side's duration using F_1 and F_2, the parameters from Model 1, and a data profile that reflects an average conflict (covariates set to a mean or median value). For each value of z_R and z_G, I simulate durations from F_1 and F_2 and record the proportion of cases where the government or the rebel group has the minimum duration.

In Figure D.2, I vary the government's initial reputation z_G from 0.90 to 0.99, while the rebel's initial reputation for resolve z_R is fixed to two values: a low value (0.90) and a high value (0.99). As mentioned above, a rebel victory refers to policy concessions from the government. The main thing to note here is that increasing an actor's reputation has the expected effect of making that actor more likely to win the conflict episode, but there is a notable difference in the sizes of these effects. Increasing the rebel group's reputation over this range results in only a modest 2.5 percentage point increase in the probability that the government backs down. In contrast, increasing the government's reputation over this range decreases the likelihood of concessions to the rebel group by about 6.5 percentage points. Changes in government reputation produce an effect that is 2.5 times stronger than changes in a rebel group's reputation, which provides some justification for past works that only focused on government reputation (e.g., Walter 2009), while also highlighting the importance of dyadic thinking in studying civil conflict (e.g., Cunningham, Gleditsch and Salehyan 2009).

Turning to Model 3 from the main text, we are most interested in whether the effect of regime type is changed once we include past conflict outcomes. The main points of interest about the results from Model 3 (Table I) are that the coefficients on democracy and recent concessions are both negative. This result on previous outcomes provides suggestive evidence that the current reputations for resolve are diminished by a recent concession. Interestingly, a rebel group challenging after a government concession is also thought less likely to be resolved, which might suggest that they are viewed as more opportunistic for challenging a government that was recently shown to be unresolved in another conflict. Overall, the negative coefficient on past concession is highly intuitive and provides some reasonable validity to the model.

Figure D.2: Two-sided reputations and conflict outcomes

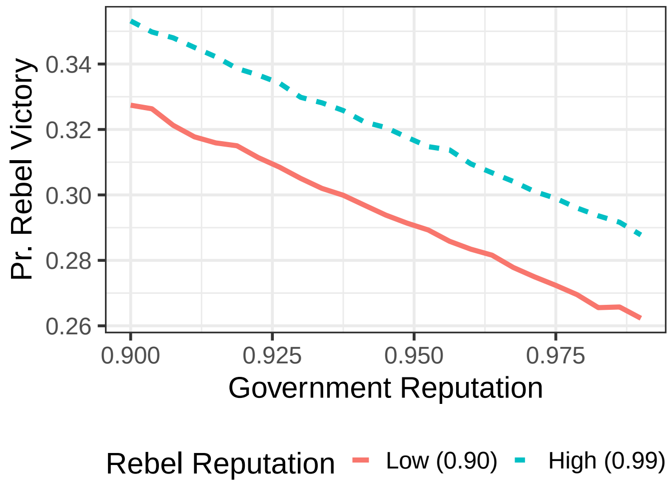
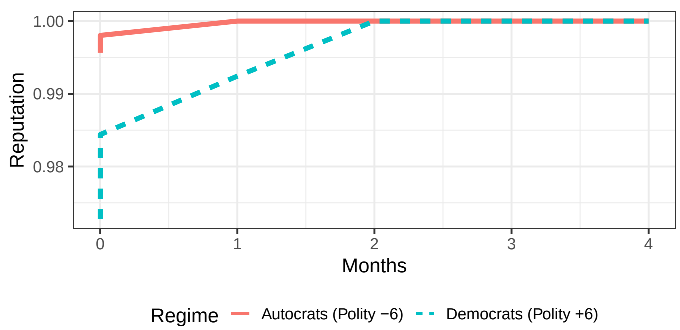


Figure D.3: Reputations for resolve across regimes (Accounting for past conflict)



We are also interested in whether the regime gap still appear in Model 3. In Figure D.3 we see that the effects mostly match those in the main text (Figure 1). Overall, while we still suspect that past concessions will negatively effect reputations for resolve, they do not appear to change the effect of democracy. Likewise, we still see the reputation jumps once conflict begins.

D.2 Different specifications of z

In this section I consider four robustness checks that use different specifications and measurements for the reputation for resolve parameters z. This set of checks provides some evidence that the main results are not driven by choices about the covariates. In all cases we look only at the estimates in z, as these are the results of interest.

Next, Table D.3 considers all the variables that were included in λ except the indicators for a coup and an oil exporting state. Attempts at including these final two indicators induced too much numeric instability. As such, I focus on adding military capacity and population. Population is often considered a very important civil conflict variable, and it is worth considering its inclusion here.

Table D.2: Results for z_R and z_G including military capacity

	Rebel	Government
Intercept	7.67	7.03
	$(5.63, 10.80)$	$(5.31, 10.58)$
Democracy	-0.27	-0.30
	$(-0.50, -0.05)$	$(-0.51, -0.13)$
GDP per capita (log)	-0.87	-1.31
	$(-1.74, 0.39)$	$(-2.00, -0.26)$
Territorial Conflict	-0.46	-0.02
	$(-1.99, 1.31)$	$(-1.42, 1.80)$
Mt. Terrain	-0.13	-0.31
	$(-0.75, 0.39)$	$(-0.84, 0.30)$
Ethnic Frac.	-4.22	-4.05
	$(-6.92, -2.16)$	$(-5.98, -2.07)$
Religious Frac.	-2.67	-3.58
	$(-4.71, -0.51)$	$(-6.53, -1.29)$
Miltary personnel per capita (log)	0.36	0.88
	$(0.05, 0.86)$	$(0.43, 1.26)$
Log L	-1532.70	
N	397	

Notes: Bootstrapped 95% confidence interval in parentheses

In the next two set of results, I return to the main model specification in Table I, but with measurement changes. The first uses alternative measures for democracy, rugged terrain, and separatist conflicts; the second recodes some of the conflict outcomes. The former uses Varieties of Democracy (V-DEM) three category measure (Pemstein, Marquardt, Tzelgov, Wang, Medzihorsky, Krusell, Miri and von Römer 2019), the (logged) standard deviation of rugged terrain from Shaver, Carter and Shawa (2019), and Stewart's (2018) measure of separatist conflicts (supplemented with the UCDP incompatibility indicator when I did not find an exact match). These results are presented in Table D.4. The latter recodes all ceasefires to indicate rebel's backing down to ensure that the main results do not hinge on how ceasefires are classified. These results are presented in Table D.5. Note that the main result holds in both cases.

We might also wonder about the intensity of conflicts. After all, we saw in the main results that both governments and rebels tend to have worse reputations for resolve within democracies. One reason we might expect this to be the case is if conflicts within democracies are just fundamentally different. It might be the case that these conflicts are just lower intensity disputes that involve more terrorism and less conventional fighting. If this is true, then maybe the observed relationship with democracy is simply an artifact of the types of disputes faced by democratic governments. However, democracy is already included in the hazard rate specification (see Table D.1), as such differences in the types of conflicts faced by democracies could be captured here if the types of conflicts involving democracies fundamentally vary in terms of goals ($\bar{\alpha}$), time horizons (r), or costs (κ).

That said, we may want to directly explore reputational differences associated with different

Table D.3: Results for z_R and z_G including military capacity and population

	Rebel	Government
Intercept	8.00	9.51
	$(4.89, 12.59)$	$(7.70, 13.21)$
Democracy	-0.27	-0.21
	$(-0.63, -0.05)$	$(-0.57, -0.05)$
GDP per capita (log)	-0.70	-1.34
	$(-1.95, 0.88)$	$(-2.16, -0.47)$
Territorial Conflict	-0.69	0.01
	$(-3.19, 1.70)$	$(-1.72, 3.11)$
Mt. Terrain	0.07	-0.29
	$(-0.79, 0.70)$	$(-1.05, 0.51)$
Ethnic Frac.	-5.78	-2.44
	$(-10.10, -3.39)$	$(-4.54, -0.74)$
Religious Frac.	-2.19	-3.13
	$(-5.17, -0.19)$	$(-5.64, -0.84)$
Miltary personnel per capita (log)	0.19	1.15
	$(-0.37, 1.86)$	$(0.63, 1.68)$
Population (log)	0.13	-0.49
	$(-0.09, 0.88)$	$(-0.81, -0.17)$
Log L	-1532.30	
N	397	

Notes: Bootstrapped 95% confidence interval in parentheses

Table D.4: Results for z_R and z_G using alternative measures

	Rebel	Government
Intercept	14.44	11.29
	$(10.64, 23.34)$	$(7.93, 21.18)$
Democracy (V-Dem)	-1.60	-4.35
	$(-8.18, 3.53)$	$(-7.55, -0.94)$
GDP per capita (log)	-0.46	-1.12
	$(-1.90, 1.56)$	$(-3.72, -0.06)$
Separatist Conflict (Stewart 2018)	0.67	1.64
	$(-1.85, 7.80)$	$(-3.73, 6.63)$
Rugged Terrain (Shaver, Carter and Shawa 2019)	0.20	0.27
	$(-0.43, 1.60)$	$(-0.58, 1.54)$
Ethnic Frac.	-14.85	-4.50
	$(-21.65, -9.71)$	$(-9.09, 0.27)$
Religious Frac.	1.84	-3.33
	$(-5.82, 7.87)$	$(-11.16, 2.14)$
Log L	-1602.13	
N	415	

Notes: Bootstrapped 95% confidence interval in parentheses

Table D.5: Results for z_R and z_G recoding ceasefires

	Rebel	Government
Intercept	6.86	7.20
	$(5.12, 21.07)$	$(5.11, 11.65)$
Democracy	-0.19	-0.20
	$(-0.96, -0.07)$	$(-0.49, -0.005)$
GDP per capita (log)	-0.33	-0.15
	$(-1.52, 1.93)$	$(-1.91, 0.99)$
Territorial Conflict	1.24	15.87
	$(-7.12, 3.72)$	$(6.95, 22.53)$
Mt. Terrain	-0.14	-0.10
	$(-1.40, 6.06)$	$(-0.89, 0.62)$
Ethnic Frac.	-0.96	-2.53
	$(-6.00, 4.23)$	$(-5.38, 1.24)$
Religious Frac.	-5.31	-3.28
	$(-12.70, 0.75)$	$(-7.15, 5.14)$
Log L	-1509.38	
N	397	

Notes: Bootstrapped 95% confidence interval in parentheses

types of conflicts. To do this, I consider a model that includes the conflict's intensity level using UCDP's measure. This variable takes on a value of 1 if the conflict ever has more than 1,000 battle deaths in a year and a 0 otherwise. This variable is post-treatment, but it does provide an important check on conflict differences. If we think that actors have a good idea for how intense a conflict is likely to get at the start of fighting, then this measure can be thought of as a proxy for this expectation. Overall, about 20% of all conflicts reach this intensity. This number jumps to 30% when we look at just autocracies and drops to about 8% in democracies. As such, it appears to capture the differences in conflicts across regime types. The results from this model are presented in Table D.6. While conflict intensity appears to have a strong effect, the overall main result on democracy is unchanged.

D.3 Democratic traits and reputation for resolve

In Table D.7 we break down the measure of democracy into three main traits: competitiveness, constraints on the use of power, and freedom of the press. The first two components are taken from the polity IV dataset, while freedom of the press data is from Li (2005) and Freedom House (Karlekar and Dunham 2012). Breaking democracy into several components allows for a more thorough exploration of the mechanisms through which democracy impacts a state's reputation for resolve. Here we see a few interesting results. First, executive constraints are associated with a boost in a state's reputation for resolve. This result is intuitive and matches various punishment-based democratic credibility theories (as in Partell and Palmer 1999). Second, both a free press and the competitiveness measure are associated with a decrease in a state's reputation for resolve.

Table D.6: Results for z_R and z_G accounting for conflict intensity

	Rebel	Government
Intercept	9.53	11.16
	$(7.71, 14.94)$	$(8.56, 16.27)$
Democracy	-0.21	-0.18
	$(-0.79, -0.01)$	$(-0.52, -0.01)$
GDP per capita (log)	-0.51	-1.44
	$(-1.88, 1.98)$	$(-2.73, 0.17)$
Territorial Conflict	-0.14	1.29
	$(-3.61, 4.83)$	$(-2.01, 6.60)$
Mt. Terrain	0.32	-0.46
	$(-0.61, 1.93)$	$(-1.42, 0.68)$
Ethnic Frac.	-6.16	-3.02
	$(-8.94, -0.03)$	$(-5.47, 1.07)$
Religious Frac.	-2.53	-5.72
	$(-5.82, 3.48)$	$(-10.28, -0.84)$
Conflict intensity	14.21	18.09
	$(3.79, 21.56)$	$(6.37, 24.22)$
Log L	-1534.68	
N	397	

Notes: Bootstrapped 95% confidence interval in parentheses

Both of these factors are consistent with volatile voter theories, wherein the competitive nature of democracy and the accountability to voters prone to war wariness. Overall, the negative effect associated with competition captures the public's ability to exert its war wariness and pressure on the executive, which offers a plausible explanation for why democracies are disadvantaged in terms of building a reputation for resolve.

The direction of these findings is consistent with the main expectations, where executive constraints are frequently viewed as a stand-in for the prospect of punishment and as such should raise a regime's reputation for resolve, while competitiveness and free press capture the public's fickleness within democratic regimes and their ability to exert this fickleness. This split finding suggests that competing mechanisms exists within democracies, but the negative effect of answering to voters and their concerns about democratic norms dominates the credibility effects of internal constraints/punishments.

D.4 Selection into violent conflict

As mentioned above, it is easy to imagine that conflicts ending at time $t = 0$ are actually conflicts that never escalate to violence. This might mean that we are under counting the number of length zero conflicts, which may potentially bias the results. One way we can address this concern is by adding in some of these missing zeros by looking at situations of non-violent conflict that did not escalate to violence. For these cases, I turn to the Conflict Information and Analysis System

Table D.7: Results for z_R and z_G with the components of democracy

	Rebel	Government
Intercept	11.67	12.21
	$(6.11, 23.65)$	$(8.43, 23.67)$
Executive constraints	0.30	0.53
	$(-2.03, 1.83)$	$(-0.22, 1.55)$
Competitiveness	-1.07	-1.17
	$(-1.91, 1.89)$	$(-2.17, -0.52)$
Free Press	4.35	-0.73
	$(-3.64, 16.55)$	$(-8.65, 6.60)$
GDP per capita (log)	0.07	-0.78
	$(-1.91, 4.45)$	$(-3.56, 1.01)$
Territorial Conflict	-0.24	0.41
	$(-7.65, 4.83)$	$(-2.31, 6.57)$
Mt. Terrain	-0.32	-0.75
	$(-1.26, 2.21)$	$(-1.92, 1.36)$
Ethnic Frac.	-5.91	-2.28
	$(-12.43, 5.97)$	$(-8.00, 2.32)$
Religious Frac.	-0.99	-4.48
	$(-11.39, 5.57)$	$(-12.35, 2.70)$
Log L	-1220.76	
N	306	

Bootstrapped 95% confidence interval in parentheses

(CONIAS) data (as discussed in Bartusevičius and Gleditsch 2018). This data includes a wide-range of violent and non-violent civil conflicts. I use two approaches to identify a set of zero-duration cases to include. First, I use cases where:

- violence is threatened by at least one side, but the dispute ends before violence begins (a non-violent dispute).

However, there are only a small number of these cases (16). As such, the second approach combines these non-violent cases with the short (less than a month cases) violent cases from the main model. This approach creates a broader set of cases where either:

- violence is threatened, but one backs down before its used; or

- violence is included in the initial threat/demand and one actor accommodates the other rather than escalating past the initial period.

We start with the cases where violence was threatened but never employed in the CONIAS data. This check introduces a set of possible zeros. Here, a time of $t = 0$ is characterized by a non-violent bargaining period where the threat of violence exists. The violent form of the war of attrition then begins when violence erupts. These results are presented in Table D.8.

Table D.8: Results for z_R and z_G when only cases non-violent are counted as length 0

	Rebel	Government
Intercept	1.39	0.57
	$(-0.31, 5.51)$	$(-0.34, 2.21)$
Democracy	-0.13	-0.07
	$(-0.41, -0.01)$	$(-0.18, 0.04)$
GDP per capita (log)	-1.13	-0.27
	$(-2.12, -0.28)$	$(-0.77, 0.43)$
Territorial Conflict	0.76	0.10
	$(-1.09, 3.40)$	$(-0.69, 1.26)$
Mt. Terrain	0.54	0.51
	$(-0.05, 2.70)$	$(0.02, 0.92)$
Ethnic Frac.	1.52	2.75
	$(-1.96, 6.10)$	$(0.98, 5.19)$
Religious Frac.	4.16	0.14
	$(1.03, 9.10)$	$(-2.81, 2.89)$
Log L	-1661.80	
N	413	

Notes: Bootstrapped 95% confidence interval in parentheses

The main thing to note in this model is that the coefficient on democracy still has a negative effect on the government's initial reputation for resolve. Additionally, we see that while the magnitude of the coefficient is smaller than the other models, it is still roughly similar. Furthermore, we note that while the 95% confidence interval now contains zero, in about 90% of the bootstrap iterations the coefficient is negative. Overall, this model still provides suggestive evidence for the main result, even though this evidence is weaker than in the other models.

We can further explore the differences between this model and the one in the main text by considering how reputation for resolve evolves when we see conflicts go from non-violent to violent (i.e., when a threat of violence is acted upon). This result is presented in Figure D.4. Again, we see the reputation gap between regimes, which again provides some additional support for the main results presented in the main text.

As mentioned, this model notably shrinks the set of 0 duration cases. This is unsurprising as it is generally more difficult to find and record cases of non-violence. Overall, the number of 0 cases is reduced by about 67% from 45 to 16. An alternative approach is to pool the 0 duration cases from this model with those in the main model. The characterization of 0 duration cases is now: cases where violence is threatened but the dispute ends before either side uses violence; OR violence is included as part of the initial demand, and one actor backs down right away rather than continuing the conflict. This approach combines the 45 original $t = 0$ cases with the 16 terminated non-violent cases. The results are presented in Table D.9.

Overall, I have considered three ways to consider what it means for a conflict to end at $t = 0$. In the main text, these are flares of violence that quickly deescalated in less than a month. This

Figure D.4: Reputations for resolve across regimes with non-violent cases

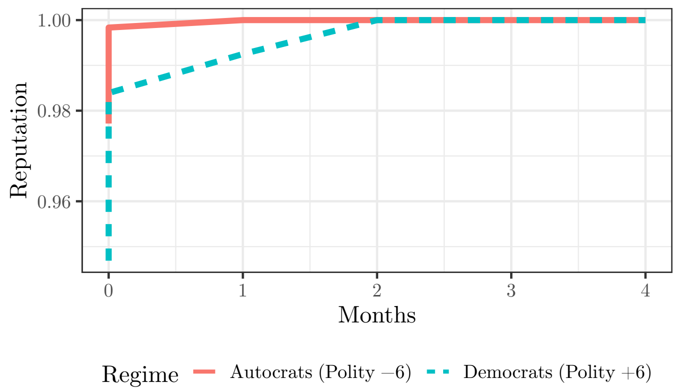


Table D.9: Results for z_R and z_G with violent and non-violent zero-duration cases

	Rebel	Government
Intercept	9.34	2.06
	$(5.80, 13.29)$	$(0.69, 3.53)$
Democracy	-0.20	-0.11
	$(-0.49, -0.005)$	$(0.22, -0.004)$
GDP per capita (log)	-1.63	-0.22
	$(-2.67, -0.50)$	$(-0.81, 0.29)$
Territorial Conflict	-2.36	-0.26
	$(-5.40, -0.26)$	$(-1.32, 0.66)$
Mt. Terrain	0.70	0.36
	$(0.22, 1.43)$	$(-0.002, 0.76)$
Ethnic Frac.	-5.63	1.95
	$(-7.58, -1.05)$	$(0.70, 3.62)$
Religious Frac.	-0.46	-1.81
	$(-3.69, 2.35)$	$(-3.67, 0.31)$
Log L	-1683.44	
N	413	

Bootstrapped 95% confidence interval in parentheses

approach characterized $t = 0$ as cases where violence began but one side immediately backed down from it. In this appendix, I first looked at a set of cases where violence was threatened, but one side backed down before the violent war of attrition begins. Finally, I combined the above approaches to fully explore how sensitive the model is to different conceptualizations. While there is some sensitivity in the magnitude and strength of the democracy result, all three approaches find a negative relationship between democracy and reputation for resolve.

As an additional consideration, I conduct a new Monte Carlo analysis using the same parameters as above, but where an unobserved variable $W_m \sim Bernoulli(0.5)$ determines whether or not conflict m is actually observed, where m indexes all the conflicts where $y_n = 0$. What this means is that for the set of conflicts with zero duration, roughly half will be unobserved to the analyst. With this large amount of missing data, we want to know how much bias (and in what direction) appears in the point estimates. At this level of missingness, the estimates on covariates of interest display some attenuation bias, which provides some confidence in that the main results are at least capture the directions of the true relationships of interest, while underestimating the magnitude.

E More details on the empirical and theoretical model

Recall that an equilibrium is characterized by a pair of CDFs F_R and F_G, where $F_i(t)$ describes the joint probability that i is unresolved and has exited the conflict at time t or sooner. These two functions have four important properties that lead to their definition. First, an unresolved player never hesitates to exit a conflict once it knows that its opponent is resolved. This first property tells us that for all t, $F_i(t) \leq 1 - z_i$, which is to say that the probability that i is both unresolved and exits the conflict by time t is at most the initial probability that i is unresolved. This tells us that either $F_i(t) < 1 - z_i$ for all $t < \infty$ or there is some $T_i < \infty$ where an unresolved i must have exited the conflict and then for all $t \geq T_i$, $F_i(t) = 1 - z_i$. Since, an unresolved player j won't hesitate to exit the conflict once it is certain that i is resolved, let $T = \min\{T_i, T_j\}$ be the point where no unresolved actors remain in the conflict.

Second, F is continuous for both sides. This follows from the fact that if F_i has a discontinuous increase at t, then j is content to pay the instantaneous cost of delay (conflict) prior to the jump in order to benefit from the discrete gain in the likelihood that i concedes at t. Third, for any t to t' interval where F_i is constant, F_j will also be constant. To see this, note that if i does not quit between t to t' (F_i is constant), then any increases in F_j on that interval make j worse off (by raising j's own likelihood of losing without receiving any gains in the likelihood that i loses) than either quitting at t or also refusing to quit on this interval. Fourth, for any interval (t, t'), where $t' < T$, neither F_i nor F_j is constant. Taken together, these four aspects all imply F_i and F_j are continuous and strictly increasing on the interval $[0, T]$.

By applying the definition of a hazard rate F_i can be rewritten as

$$F_i(t) = 1 - (1 - F_i(0))e^{-\lambda_i t}.$$

Recall that at time T, $F_i(T) = 1 - z_i$, which gives us

$$F_i(T) = 1 - (1 - F_i(0))e^{-\lambda_i T} = 1 - z_i,$$

and solving for T we find that

$$T = -\frac{1}{\lambda_i} \log\left(\frac{z_i}{1 - F_i(0)}\right).$$

Since we $F_2(0) = 0$, it can be shown that an actor is the possible first quitter only if $\lambda_1^{-1} \log(z_1) < \lambda_2^{-1} \log(z_2)$, and that for this player $F_1(0) = 1 - z_1 z_2^{-\lambda_1/\lambda_2}$. After this first moment actors exit at a constant rate, leading to the strategies presented in the text. Where the complete CDF's are

$$F_1(t|\lambda, z) = \begin{cases} 1 - z_1 & t \geq T \\ 1 - z_1 z_2^{-\lambda_1/\lambda_2} e^{-\lambda_1 t} & 0 \leq t < T \end{cases}$$

and

$$F_2(t|\lambda) = \begin{cases} 1 - z_2 & t \geq T \\ 1 - e^{-\lambda_2 t} & 0 \leq t < T. \end{cases}$$

where G is the possible first quitter when $\lambda_G^{-1} \log(z_G) < \lambda_R^{-1} \log(z_R)$ as before.

Setting $F_1(0) = 1 - z_1 z_2^{-\lambda_1/\lambda_2}$ makes T the same for both actors and introduces the constraint that

$$-\frac{1}{\lambda_1} \log\left(\frac{z_1}{1 - F_1(0)}\right) = -\frac{1}{\lambda_2} \log\left(z_2\right).$$

Note that this constraint is not included in the estimation. This omission done to avoid zero and non-monotonic likelihood problems and makes the estimation procedure a pseudo-likelihood approach. A pseudo-likelihood routine relaxes a dependency in the data generating process in order to gain tractability and suggests that the final estimates may not be in equilibrium, but they are informed by it through the remaining conditions of the model. Two post-estimation diagnostics give us additional confidence in the approach. First, note that in Figure D.1, the actors' beliefs both reach 1 at nearly the same time, suggesting that the constraint implied by T is informing the estimation procedure indirectly through the identification of the possible first quitter. Indeed in Appendix F, I show that the constraint implied by T holds at the point estimates in nearly 70% of observations. Second, the other model fit statistics in Appendix F suggest that the model still performs relatively well when applied to data.

E.1 Constant costs and hazards

Recall that λ_i reflects a constant hazard rate at which i exits the conflict and part of that hazard rate is j's constant cost of fighting κ_j. There are of course good reasons to suspect that neither of these assumptions is particularly realistic in civil conflict. As battlefield conditions change or finances improve/worsen for groups, it is easy imagine that costs, and as a result the hazard rate, should adjust to reflect these new conditions. This raises two important questions for the analysis: Do we need constant costs/hazard rates and is the empirical model fit affected by this assumption?

The answer to the first question is yes. Past work as shown that a large class of war-of-attrition models, including the one presented here, require constant hazard rates in order for strategies to be in equilibrium (Abreu and Gul 2000; Hendricks, Weiss and Wilson 1988). As a result, a theoretical model that captures non-constant hazards/costs requires building a new theoretical framework for understanding bargaining and reputation. Future work should absolutely consider such an enterprise, as it would undoubtedly have merit in terms of empirical realism. However at this time the theoretical tools for understanding bargaining with reputation concerns demonstrate that any equilibrium to the continuous time bargaining game will involve constant hazards.

The second question is addressed in Appendix F, below, where I find that the strategic model with constant hazards fits the data roughly as well as non-strategic models with more flexible hazard rates (e.g., Cox proportional hazard model). Recent work by Dion, Boehmke, MacMillian and Shipan (2016) has considered an empirical war of attrition that allows for non-constant hazards. However, it is unclear how introducing this flexibility would affect the interpretation of this model as it introduces a gap between the theoretical and empirical models. This additional gap would makes it increasingly problematic to interpret λ and z within the structure of the formal model as the constant hazard is a primary feature of the war of attrition. As in all structural work, the ability to interpret the estimates within the model structure is a key identifying assumption.

F Model fit

Some model fit diagnostics using Model 2 are reported in Table F.1. In the first row of Table F.1, I consider the models using Cox-Snell's generalized R^2. This measure shows the general improvement in fit for each model over a constant only version of the same model. These are not directly comparable across models, but they provide a sense as to how much the same covariates improve each individual model. Higher values here are good; the strategic model's large generalized R^2 provides us with some confidence in the results.

The remaining statistics are comparative in nature. I compare the three models by calculating the root mean squared error (RMSE) on the in-sample predicted duration (only for conflicts that have ended).[6] Overall, we see that these three models all perform roughly the same in this dimension, with the strategic model performing slightly better than the rest.

Table F.1: Model fit statistics

	Strategic Duration	Ordinary Weibull	Ordinary Cox	Ordinary Logit
Cox-Snell R^2	0.48	0.04	0.03	–
RMSE: In-sample	53.95	55.69	55.76	–
RMSE: Out-of-sample	33.40	50.92	37.42	–
Expectation: % Correctly Predicted	0.44	–	–	0.44

The next thing I consider is out-of-sample fit. In this case, I use 37 civil wars that appear in

[6]To produce predicted duration from the Cox model, I estimate the baseline hazard using Breslow's method.

Fearon's (2004) dataset but not in the data used to fit the models. I code the victors of these cases following the UCDP codebook, and I employ conflict encyclopedias to determine the duration in months. The strategic estimator convincingly wins this contest. Out-of-sample prediction is usually considered an excellent test of model fit; these results give an edge to the strategic model.

The final quantity reported in Table F.1 is the percentage of government concessions correctly predicted (in-sample). In this case, we compare the strategic duration model's ability to produce probabilities of rebel victories to an ordinary logit with the sample specification. Given that the logit is the standard tool for binary choice modeling and the strategic duration model only produces these estimates as a side effect, we might expect the logit to dominate here. Instead we find that the two are virtually tied.

A separate model fit statistic concerns only the pseudo-likelihood estimator. Specifically, we want to know if the omitted equilibrium constraint associated with T is still satisfied at the estimates (see Appendix E for more details). Because the constraint is not imposed during estimation there is no guarantee that it will be satisfied at any observation. We want to know how badly the constraint is violated. Larger violations could indicate poor fit. Using the definition of T in Appendix E, we find that more than two-thirds of observations the omitted equilibrium condition is still satisfied within a tolerance of 1×10^{-5}. Additionally, the mean violation is only about 1.5. Finding that this condition holds reasonably provides us with some relief in omitting it from the estimation procedure.

References

Abreu, Dilip and Faruk Gul. 2000. "Bargaining and Reputation." *Econometrica* 68(1):85–117.

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

Chernozhukov, Victor and Han Hong. 2004. "Likelihood Estimation and Inference in a Class of Nonregular Econometric Models." *Econometrica* 72(5):1445–1480.

Crisman-Cox, Casey and Michael Gibilisco. 2018. "Audience Costs and the Dynamics of War and Peace." *American Journal of Political Science* 62(3):566–580.

Cunningham, David E., Kristian Skrede Gleditsch and Idean Salehyan. 2009. "It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53(4):570–597.

Dion, Douglas, Frederick J. Boehmke, William MacMillian and Charles R Shipan. 2016. "The Filibuster as a War of Attrition." *The Journal of Law and Economics* 59(3):569–595.

Fearon, James D. 2004. "Why do some Civil Wars Last so much Longer than Others?" *Journal of Peace Research* 41(3):275–301.

Hendricks, Ken, Andrew Weiss and Charles Wilson. 1988. "The War of Attrition in Continuous Time with Complete Information." *International Economic Review* 29(4):663–680.

Karlekar, Karin Deutsch and Jennifer Dunham. 2012. "Freedom of the Press 2012." Freedom House. Published at: http://www.freedomhouse.org/report-types/freedom-press.

Kreutz, Joakim. 2010. "How and When Armed Conflicts End: Introducing the UCDP Conflict Termination Dataset." *Journal of Peace Research* 47(2):243–250.

Li, Quan. 2005. "Does Democracy Promote or Reduce Transnational Terrorist Incidents?" *Journal of Conflict Resolution* 49(2):278–297.

Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.

Mullen, Katharine, David Ardia, David Gil, Donald Windover and James Cline. 2011. "DEoptim: An R Package for Global Optimization by Differential Evolution." *Journal of Statistical Software* 40(6):1–26.
URL: *http://www.jstatsoft.org/v40/i06/*

Partell, Peter J. and Glenn Palmer. 1999. "Audience Costs and Interstate Crises: An Emprical Assessment of Fearon's Model of Dispute Outcomes." *International Studies Quarterly* 43(2):389–405.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri and Johannes von Römer. 2019. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." V-Dem Working Paper No. 21. 4th edition. University of Gothenburg: Varieties of Democracy Institute.

Ramsay, Kristopher W. and Curtis S. Signorino. 2009. "A Statistical Model of the Ultimatum Game." Unpublished Manuscript.

Shaver, Andrew, David B. Carter and Tsering Wangyal Shawa. 2019. "Terrain ruggedness and land cover: Improved data for most research designs." *Conflict Management and Peace Science* 36(2):191–218.

Stewart, Megan A. 2018. "Civil War as State-Making: Strategic Governance in Civil War." *International Organization* 72(1):205âĂŞ226.

Walter, Barbara F. 2009. *Reputation and civil war: why separatist conflicts are so violent*. New York: Cambridge University Press.

Whang, Taehee, Elena V. McLean and Douglas W. Kuberski. 2013. "Coercion, Information, and the Success of Sanction Threats." *American Journal of Political Science* 57(1):65–81.