Rage Against the Machines and the Democratic Hourglass?

A Replication Project

Brian Forester

27 April 2012

Abstract

In counterinsurgency wars, what accounts for the decreasing likelihood of insurgent defeat over the course of the twentieth century? Lyall and Wilson answer with the increasing mechanization of counterinsurgent military forces.[1] Using a newly constructed data set of 286 counterinsurgencies, the authors test their argument with ordered logistic regression models, finding statistical support for their mechanization thesis. This article builds off of Lyall and Wilson's study by first replicating their findings, then conducting a series of simulated experiments to further clarify the results, and finally proposing a model re-specification to capture the interactive effects of conflict duration and regime type on counterinsurgent victory.

Data and Methods

The central question undergirding this study is best depicted by figure 1. Counterinsurgency outcomes have become less favorable to the counterinsurgent (incumbent) since the middle of the nineteenth century. The unit of analysis for this study is counterinsurgency war.

The authors define an insurgency as "a protracted violent struggle by nonstate actors to obtain their political objectives - often independence, greater autonomy, or subversion of existing authorities - against the current political authority"(70). Additionally, each case must meet a threshold of 1,000 battle deaths, with at least 100 casualties on each side, and the adoption of a guerilla warfare strategy by the nonstate actor (insurgent). The 286 cases meeting these definitional requirements were drawn from four separate data sets.[2]

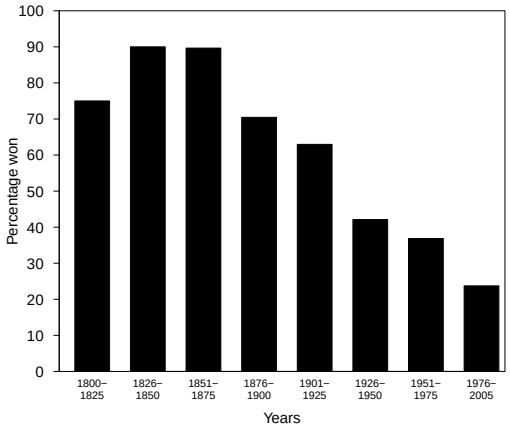


Figure 1: *The figure depicts the percentage of conflicts won by incumbents over time. There are 286 cases from 1800 to 2005.*

[1] Jason Lyall and Isaiah Wilson. "Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars." *International Organization*, 63 , pp 67-106, Winter 2009.

[2] The Correlates of War (3.0); the Fearon-Laitin civil war data set (2003); the Uppsala Armed Conflict Dataset (version 3); and the Political Instability Task Force's (PITF) Internal Wars Dataset

The dependent variable in this study is the outcome of the conflict, and it is coded from the perspective of the incumbent. In the tradition of previous conflict studies, the authors operationalize the dependent variable as holding one of three ordinal categories: win, draw, or loss. Cases are assigned a 0 for a loss, a 1 for a draw, and a 2 for a win.

The principal independent variable of this investigation is the incumbent's military force structure. The authors operationalize this variable in three ways. First, they create a dichotomous variable, *modern*, that differentiates between the foraging era (1800-1917) and the machine era (1918-2005). Second, within the machine era (1918-2005), the authors measure an incumbent's level of mechanization by measuring the number of soldiers per mechanized vehicle. Using this scaled index, the authors create a fourfold ordinal variable, *mech*, with cut points at the 25 percent quartiles to represent increasing levels of mechanization. There are 167 observations of the variable *mech*. Figure 2 is a bivariate scatterplot of mechanization over time, which, when compared with figure 1, indicates the correlation that the authors are wishing to further analyze. Finally, the authors employ a dichotomous indicator, *heli*, to represent incumbents that employ 25 or more helicopters in wars during the post-1945 era. There are 135 observations of *heli*.

In addition to the key explanatory variables, competing explanations are tested and included in the statistical models. In exploring the relationship between regime type and war outcome, the authors include the variable *regime*, as measured by Polity2 values from the PolityIV data set. Additionally, the authors include measures of regime sensitivity to international pressure (*trade*) and of relative economic and military power (*power*). Also included in the analysis is an ordinal variable, *support*, indicating the level of external or internal support provided to the insurgency in each case. Finally, the authors include a dichotomous indicator, *occupy*, to represent incumbents that are external occupiers. The descriptive statistics and full definition of each independent variable and control variable used in this study can be found in Appendix 1.

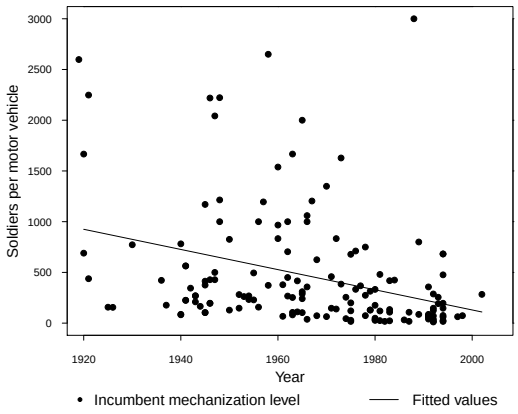


Figure 2: *This figure represents the number of soldiers per vehicle for 167 post-1918 incumbents regressed on time. The general negative trend indicates increasing mechanization of incumbent military forces over time.*

In order to test their mechanization thesis, the authors employ a statistical analysis of 286 cases of counterinsurgency. Because of the ordinal nature of the dependent variable, ordered logistic regression is used to examine general associations between the dependent and independent variables. The authors employ a two step method in testing their mechanization hypothesis. First, they use ordered logistic regression to model the variance of incumbent victory across the foraging (pre-World War I) and machine eras (post-World War I). The authors choose the first World War to temporally divide the data with the dichotomous indicator, *modern*, because scholarly convention holds that this is when mechanized militaries first began to emerge. After examining these two distinct periods, the authors then use ordered logistic regression to model the variance of incumbent victory across differing levels of mechanization within the modern era.

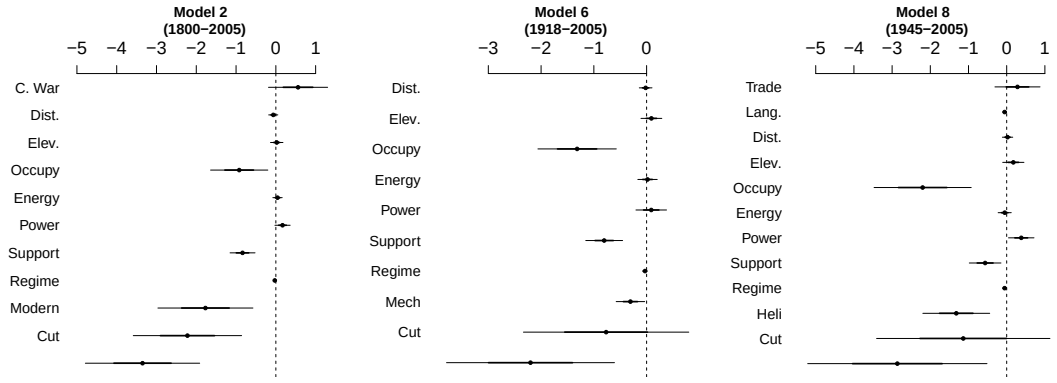
After first replicating the original findings of the authors, I will then turn to an assessment of the models' goodness of fit. Further, I will use statistical simulation to clarify our understanding of general conclusions drawn from the original results, and finally, I will propose and test a model re-specification that incorporates a new variable, *duration*, into the model.

Results

The principal findings presented by the authors are reflected in figure 3. These coefficient plots depict the estimators for each of three models that are essential to the authors' claim.[3] I will base the remainder of my analysis and extension around these three models.

In looking first at model 2, we see that the parameter estimate for *modern* is negative and "statistically significant" (i.e. statistically distinct from 0 at the 0.05 level of significance). Substantively, this means that a shift from the foraging to modern areas is generally associated with a decreased probability of incumbent victory. In fact, model 2 predicts an approximately 80% probability of incumbent victory in the foraging era and a 42% probability of incumbent victory in the machine era.[4] We can also see that the variables *occupy* and *support* appear to be strong negative predictors of increases in the dependent variable.

In model 6, the authors introduce the variable *mech*, and as the figure reflects, its estimator is negatively associated with increases in the dependent variable. Substantively, the model predicts a 52% probability of incumbent victory with a low mechanized military, and a 31% probability of incumbent victory with a highly mechanized military, when all other variables are held constant. As with model 2, the variables *occupy* and *support* are significant predictors of the war's outcome.



(a) *Modern; N=286* (b) *Mechanization; N=167* (c) *Helicopter; N=135*

Figure 3: *The above figure contains the coefficient plots of the principal ordered logistic regression results from the authors' analysis, which models the variance of conflict outcome across foraging versus modern eras (a), levels of mechanization (b), and the use of helicopters by the incumbent (c). The dots represent point estimates, and the solid line represents the 95% confidence interval.*

[3]The complete, replicated regression tables for all eight models can be found in Appendix 2.

[4]These (and all further) predicted probabilities of incumbent victory were computed with an inverse logit transformation using different values of the independent variable of interest (*modern* here) while holding all other continuous variables in the model at their means, and all dichotomous variables at their median values. Missing from these predicted probabilities is the associated uncertainty, which I will address in a later section.

Model 8 extends the authors' argument by introducing a dichotomous variable indicating incumbent use of helicopters in counterinsurgency operations. The estimated parameter for the variable *heli* demonstrates that the use of helicopters by incumbents is generally associated with a lower probability of victory. It is noteworthy, however, that the authors' operationalization of *heli* does not differentiate between use of helicopters in a logistical support role versus as a close combat attack role. Such differentiation may influence the authors' results. Substantively, though, the model predicts a 47% probability of victory for incumbents not employing helicopters and a 20% probability of victory for incumbents that do use helicopters in counterinsurgency operations. Like the previous models, the estimated coefficients for the variables *occupy* and *support* remain negatively associated with the dependent variable. The *regime* variable reaches significance only in model 8, which suggests that as states become more democratic, they become increasingly ineffective at waging counterinsurgency wars.

In order to provide a visual representation of mechanization's effects on counterinsurgency outcomes, the authors construct the figure below, which represents the conditional effects of mechanization. Using model 6 to compute predicted probabilities for the post-1918 observations, the plotted curve indicates a decreasing probability of victory as the incumbent's mechanization level increases.

Having replicated the key findings presented by the authors, I turn in the next section to an assessment of the three models discussed above. Specifically, I wish to (1) assess whether the model estimations are unduly influenced by any of the data and (2) assess the goodness of fit and predictive accuracy for each model.

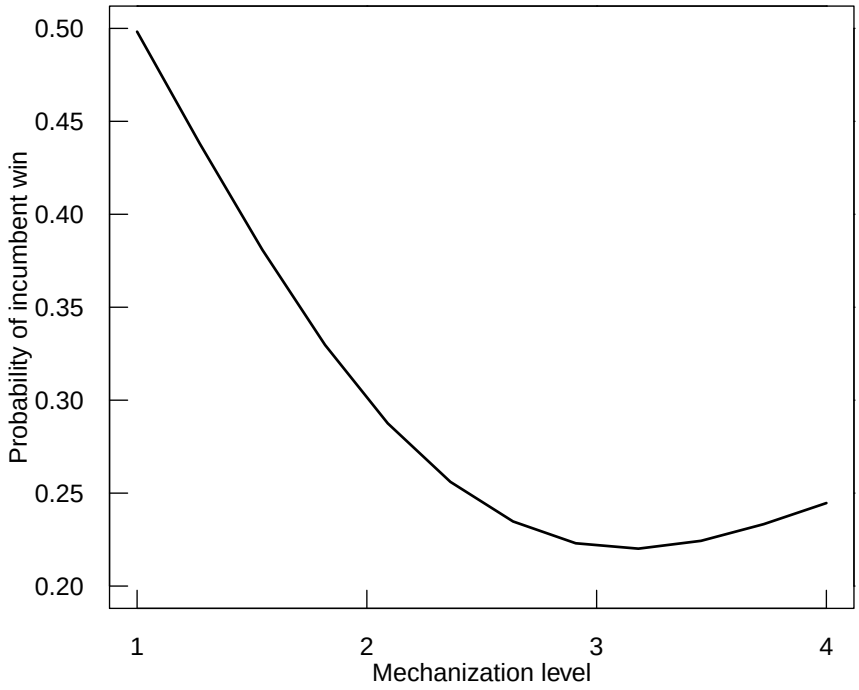
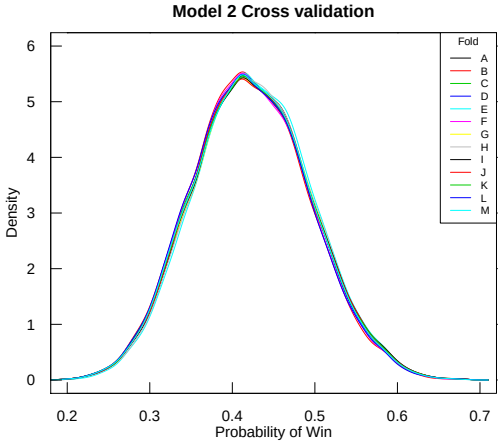


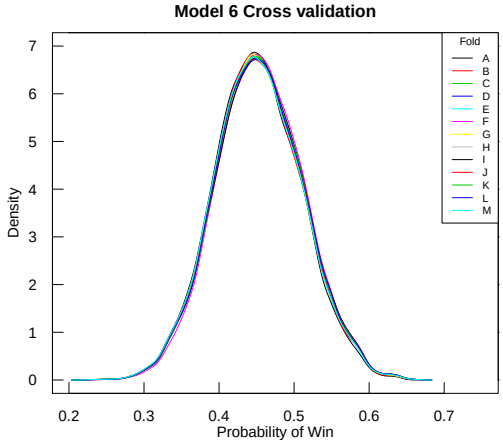
Figure 4: *This is the replication of figure 3 in the authors' analysis, which is the conditional effect of mechanization on the probability of incumbent win. The predicted probability estimates were obtained from model 6 and fitted with natural cubic spline curves. N=167.*

Model Assessments

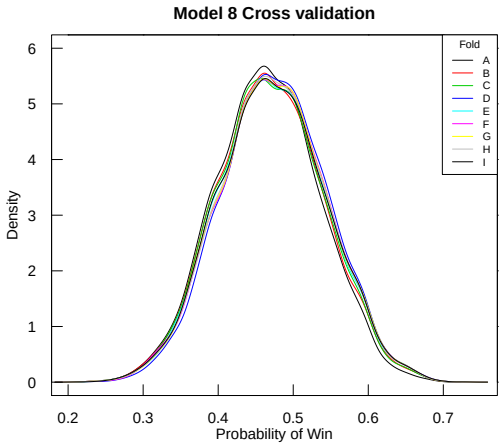
In order to comprehensively assess the models, I will first undertake a cross-validation of models 2, 6, and 8. To conduct such a test on a model, I randomly partition the respective data sets used in the estimation into roughly equivalent folds. I then leave aside one fold in the estimation of the model, simulate uncertainty through random draws of the estimates from a multivariate normal distribution, compute the predicted probability of victory while holding all variables at measures of central tendency, and plot the resulting densities. I repeat this procedure for each fold, and then each model.



(a) *Model 2; 13 fold. N=286*



(b) *Model 6; 13 fold. N=167*



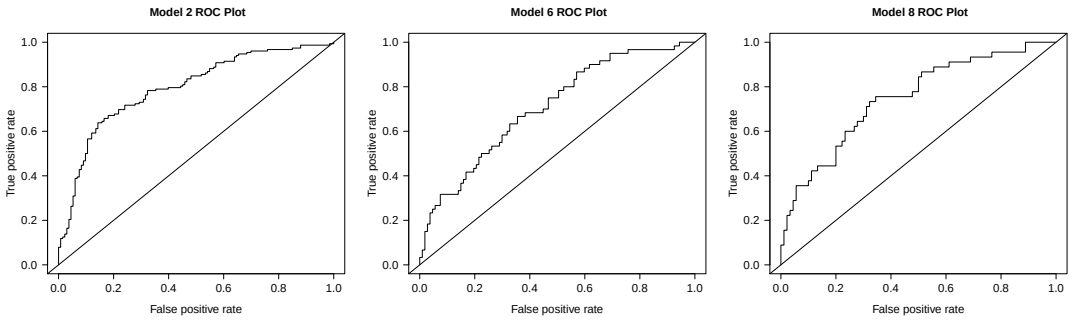
(c) *Model 8; 9 fold. N=135*

Figure 5: *The three plots above show the results of a cross-validation for each model. In computing the probability of incumbent victory, all continuous variables were held at their mean values, and all discrete variables were held at median value. The separate folds are represented by distinct colors, and coded alphabetically.*

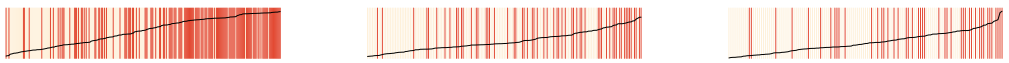
If the estimated models are not unduly influenced by any of the data, then we should expect the distributions in figure 5 to look nearly identical. If a particular fold of the data were biasing our estimates, then we should expect to see an outlying distribution. We see, though, that this is not the case across all three plots in figure 5; the data do not appear to be unduly influencing the models.

Moving from an assessment of data effects, I now wish to examine the overall predictive power of each of the models. In other words, I wish to see how accurately the models predict the outcome variable for each case, given the corresponding independent variables. More specifically, I am especially interested in how accurate the models predict incumbent victory. Therefore, in assessing the predictive power, I treat the dependent variable as a binary outcome - win (1)/no-win(0), and then examine how well the models predict incumbent wins.

In order to visualize this predictive power, I create two indicators of goodness of fit for each model. The first is the receiver operating characteristic plot (ROC). These plots illustrate the false positive and true positive rates for the model; a model that perfectly predicts outcomes would have a false positive rate of 0 across all true positive rates, and a true positive rate of 1 across all false positive rates. In the plots below, the $y = x$ line is given as a reference, and while none of the model fits are great, model 2 appears the best. The second indicator of goodness of fit depicted below is the separation plot. This simple tool orders the predicted probabilities of incumbent victory, color codes according to the actual outcome (dark=win; light=no win), and then adds a line indicating the value of the predicted probability. Of the three, model 2 appears to have the best fit. Having assessed each of the models, I next move to a deeper analysis of the independent variables of interest.



(a) *Model 2 ROC plot* (b) *Model 6 ROC plot* (c) *Model 8 ROC plot*



(d) *Model 2 separation plot* (e) *Model 6 separation plot* (f) *Model 8 separation plot*

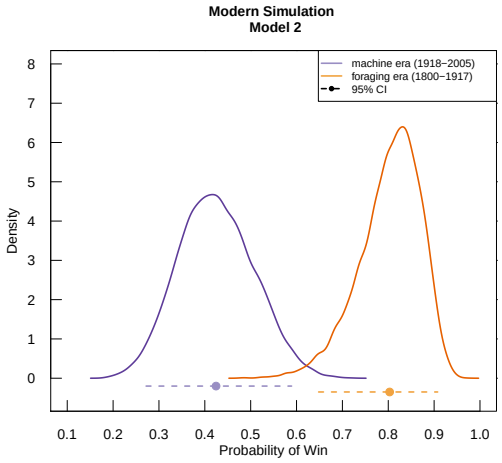
Figure 6: *The above figure contains the ROC and separation plot for each model. The ROC plots compare the true and false positive rates for each model, and the separation plots sequence the predicted probabilities, color-code the actual outcomes, and include an indicator line for the probabilities.*

Statistical Simulations

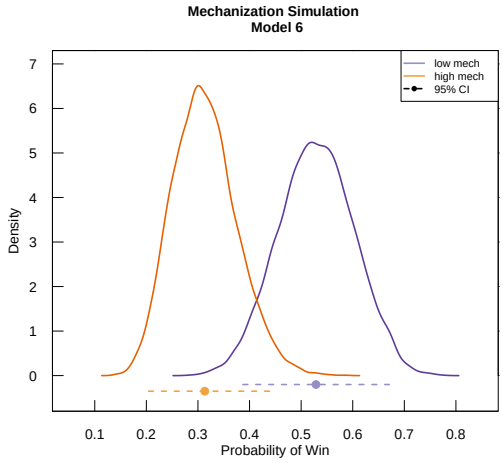
In order to further clarify our understanding of the relationship between counterinsurgency outcomes and the independent variables of interest, I will undertake a series of statistical simulations. These simulations build off of the substantive interpretations of the independent variables discussed earlier by incorporating the fundamental uncertainty inherent in our estimated models. In so doing, we are better able to visualize the true substantive meaning of our independent variables of interest.

I construct my simulation experiments in four steps. First, I create scenarios of substantive interest, where I vary the values of a particular independent variable and hold all others constant at measures of central tendency.[5] Second, I conduct 10,000 random draws of model parameter estimates from a multivariate normal distribution, thus simulating the systematic and stochastic uncertainty in the model. Third, I apply each draw to the scenarios of interest, using the inverse logit transformation to compute predicted probabilities of victory. And finally, I plot the resulting distributions.

In the figures that follow, I have plotted the results of six statistical simulations. By incorporating model uncertainty, each of these represents an opportunity to further clarify our understanding of the independent variable of interest. In figure (a), we see the effect of shifting from the foraging era to the modern era on the probability of incumbent victory, when all else is held constant. Of note here, the 95% confidence bands do not overlap, indicating a more distinctive association between differing values of the variable *modern*. Figure (b) illustrates the effect of shifting from the lowest to the highest level of incumbent mechanization on the probability of incumbent victory. In contrast to figure (a), the confidence bands do overlap, thus diminishing the certainty with which we can differentiate these two levels of mechanization. Though this result does not negate the authors' finding with respect to mechanization, it does indicate that perhaps mechanization is not as strong an effect as inferred by only looking at point estimates.

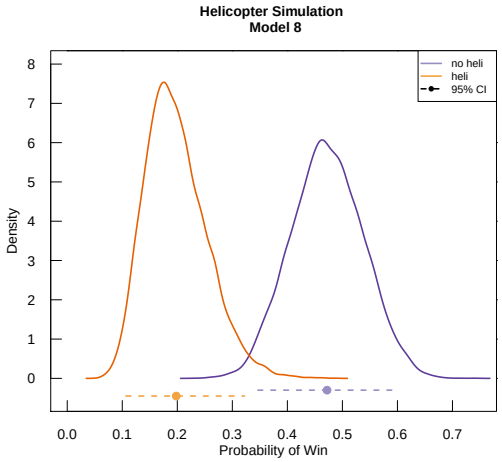


(a) *Foraging vs. Machine era*

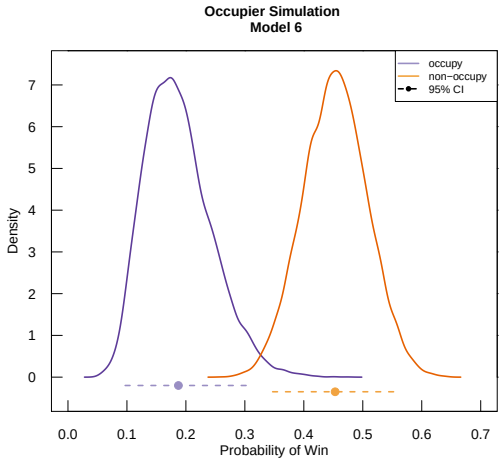


(b) *High vs. Low Mechanization*

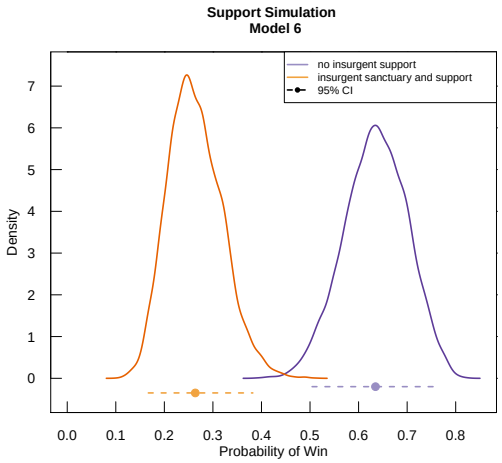
[5] As before, continuous variables at mean values and discrete variables at median values.



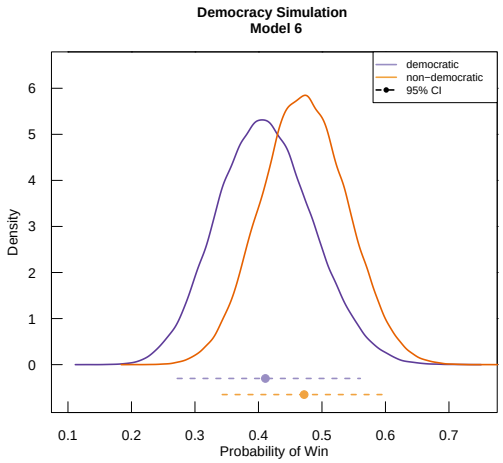
(c) *Helicopter vs. no Helicopter*



(d) *Occupier vs. non-Occupier*



(e) *No Support vs. High Support*



(f) *Democracy vs. non-Democracy*

Figure 6: *The figures above illustrate the results of various statistical simulations using the model listed in the plot title. The solid density curves are the result of 10,000 simulations. The solid circles represent point estimates, and the dotted lines represent 95% confidence bands around those estimates.*

The effect of helicopters on the probability of incumbent victory is significant, as illustrated in figure (c). A similarly strong effect is seen between occupiers and non-occupiers. Incumbents of the former type having a much higher probability of victory than the latter type. Figure (e) compares insurgencies with high levels of external support and sanctuary to insurgencies with no external support. As the distributions demonstrate, the probability of incumbent victory is significantly higher for insurgencies with no support. In order to compare democracies with non-democracies, I collapse the variable *regime* into a dichotomous indicator, with a 1 representing polity scores greater than or equal to 7, and a 0 for all others. As the results of the simulation demonstrate, there is no real discernible difference between

stable incumbent democracies and incumbent non-democracies on the predicted probability of victory.

The results of the simulations are thus illuminating. In particular, these results have clarified our understanding of mechanization's effects; it is not as strong a predictor of the outcome as simply looking at the point estimated probabilities indicates. At the same time, however, we have also demonstrated the power of *modern* and *heli* in predicting incumbent victory. In the next section, I propose a re-specification of model 6 by including a new variable, *duration*, interacting with regime type. I test this re-specified model and illustrate the results.

Model Re-Specification

Absent from the models presented in the authors' original findings is any type of variable accounting for duration of the conflict. We may expect conflict duration to have a predictive effect on the probability of incumbent victory. Furthermore, we may expect that conflict duration affects differing regime types in different ways; specifically, I expect conflict duration to affect democracies differently than non-democracies. As the length of a counterinsurgency conflict increases, domestic political will decreases in incumbent democracies, and with it, the probability of incumbent victory decreases as well. This is the argument I wish to test by a re-specification of model 6.

In order to test this re-specification, I first create a new variable, *duration*, which is operationalized by the length of the conflict from start to finish, as measured in logged days. I then interact this variable with a dichotomous indicator for democracy, *democ*, resulting in the following model specification:

$$\begin{aligned} outcome = {} & \beta_1 mech + \beta_2 democ + \beta_3 support \\ & + \beta_3 power + \beta_3 energy + \beta_4 elevation \\ & + \beta_5 distance + \beta_6 duration + \\ & + \beta_7 democ * duration + \epsilon \end{aligned} \tag{6.1}$$

This re-specified model will be estimated using the same post-1918 data used in the estimation of model 6.[6]

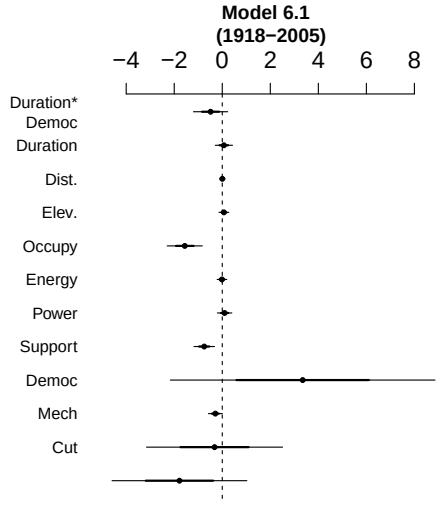


Figure 7: *The above figure represents the estimated parameters for model 6.1. The dots represent point estimates, and the solid line represents the 95% confidence interval. N=164.*

[6]Duration data was not available on 3 observations, resulting in an N of 164.

In the resulting estimation, figure 7 demonstrates that *mech*, *occupy*, and *support* remain statistically significant predictors of the dependent variable, all in the negative direction. The variable *duration*, however, does not achieve significance, nor does the interaction term. Instead of concluding no interactive effect, though, I will continue with the analysis of the re-specified model and include appropriate measures of uncertainty.

Assessments of the model fit are illustrated in figure 8. The ROC and separation plots indicate that, like model 6, the re-specified model does not do a very good job at predicting actual outcomes. Our interest, though, is the comparison of model 6 and model 6.1, and in order to meaningfully compare the two models, we can conduct a simple likelihood ratio test. If we take model 6 as our null hypothesis and model 6.1 as the alternative, then we can define the likelihood ratio test statistic, R, as follows:

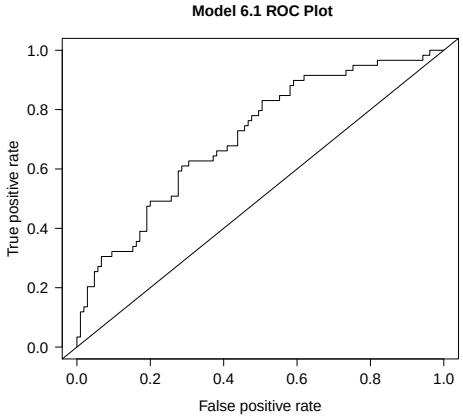
$$R = -2ln(\frac{L_0}{L_A})$$
$$R = 2(lnL_A - lnL_0)$$
$$\sim f_\chi(R|m)$$

where L_0 is the log-likelihood under the null model (6), L_A is the log-likelihood under the alternative model (6.1), and m is the difference in degrees of freedom between the two models. Substituting in the log-likelihoods from our two models:

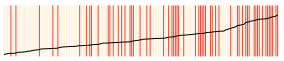
$$R = 2(-159.76 - (-162.96))$$
$$R = 6.4$$

When computing the cumulative probability of obtaining a $R = 6.4$ in a chi-square distribution with $m = 2$ degrees of freedom, we find that $p < 0.05$. Thus, using classical hypothesis testing, we can safely conclude that model 6.1 is better specified than model 6.

Having assessed the re-specified model, I now turn to an evaluation of its substantive effect on the probability of incumbent victory. In so doing, I aim to determine the differences in *duration's* effect on democracies versus non-democracies.



(a) *ROC plot*



(b) *Separation plot*

Figure 8: *This figure illustrates the goodness of fit for model 6.1. The ROC plots compare the true and false positive rates for each model, and the separation plots sequence the predicted probabilities, color-code the actual outcomes, and include an indicator line for the probabilities.*

In order to demonstrate the disparate effects of conflict duration on probability of incumbent victory, I plot the marginal effects of *duration* using the estimates from model 6.1. I first create two scenarios from the model's explanatory variables, where only the value of *democ* is different. I then conduct 10,000 random draws of the coefficients from a multivariate normal distribution, and I apply these random draws to the full range of *duration* values, which in effect creates a predicted probability distribution at each value of *duration*. The means of these distributions are connected with the solid lines, and the shaded areas represent the 95% confidence bands.

The plot illustrates the disparity in probability of winning as conflict duration increases for incumbent democracies and incumbent non-democracies; democracies' probability of victory trends negatively as *duration* increases, whereas non-democracies' probability of victory has a slight positive trend.

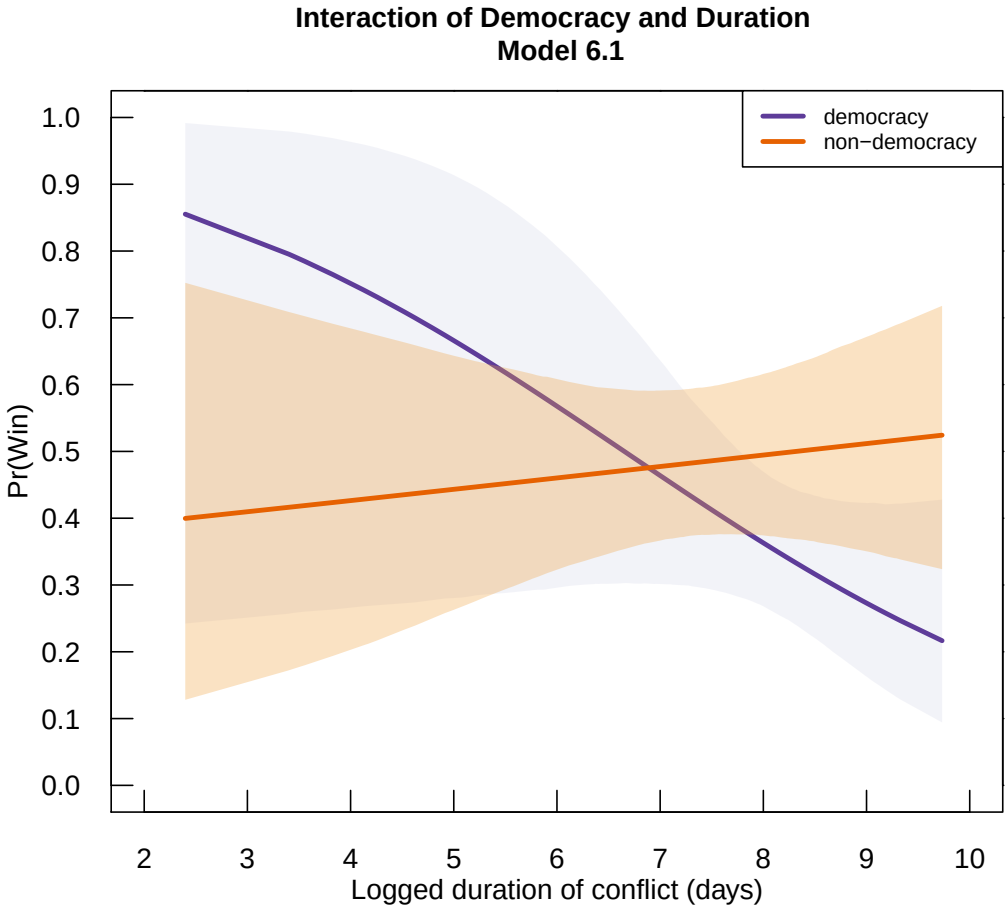


Figure 9: *This plot represents the marginal effect of conflict duration on probability of incumbent victory for democracies and non-democracies. The solid lines represent the point estimates, and the shaded areas represent the 95% confidence bands.*

To be sure, the uncertainty surrounding each of the plots makes it difficult to draw any definitive conclusions, especially for conflicts of shorter duration. However, we can clearly see the shifting of the predicted probability distributions as conflict duration increases, which is in accordance with our expectations. From a policy perspective, these findings demonstrate that democratic incumbents should be wary of engaging in potentially protracted counterinsurgency campaigns. This analysis, though, does not parse out the underlying causal mechanisms at work on democracies and non-democracies. Having identified this trend, though, serves as starting point for future research into the relationship between counterinsurgency outcomes and regime type.

References

King, Gary. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* Cambridge University Press, 1998.

Lyall, Jason and Isaiah Wilson. "Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars." *International Organization*, 63 , pp 67-106, Winter 2009.

Appendix 1: Summary statistics

	N	Std. dev.	Median	Mean	Min	Max	Definition
Outcome	286	0.86	2.00	1.26	0	2	Dependent variable: 2 if incumbent won; 1 if draw; 0 if incumbent lost.
Modern	286	0.49	1.00	0.60	0	1	Binary variable: 1 if mechanized era (1917 to 2005); 0 if not.
Railway	118	0.39	0.00	0.18	0	1	Binary variable: 1 if incumbent used railway in COIN role during foraging era (1800 to 1917); 0 if not.
Mech	167	1.13	2.00	2.45	1	4	Four values: 4 represents highest mechanization level; 1 lowest.
Heli	135	0.43	0.00	0.24	0	1	Equals 1 if incumbent used a substantial number of helicopters (25) on the battlefield; 0 otherwise.
Regime	286	7.00	0.00	-0.08	-10	10	Polity2 score (-10 to 10 scale) measured in last prewar year.
Democ	286	0.46	0.00	0.29	0	1	Equals 1 if Polity2 greater than 6, 0 if otherwise.
Trade	135	1.20	-3.41	-3.67	-9.28	-0.46	Exports + Imports as share of GDP (logged), 1946 to 2005.
Support	286	0.78	0.00	0.62	0	2	Equals 2 if sanctuary and external support; 1 if only sanctuary or support; 0 if none.
Power	286	2.25	1.15	0.53	-4.76	4.62	Natural log of an incumbents share of cumulative index of national capabilities in last prewar year.
Energy	286	3.15	-0.61	-1.55	-12.84	2.78	Natural log of incumbent energy use divided by total population in last prewar year.
Occupy	286	0.49	0.00	0.40	0	1	Equals 1 if the incumbent is an external occupier; 0 otherwise.
Elevation	286	1.38	6.21	5.98	0	8.50	Natural log of average of five elevation readings in conflict area (meters)
Distance	286	3.13	7.22	6.15	0	9.84	Natural log of distance (in kilometers) from incumbents capital to conflict area.
Language	135	7.39	4.00	8.04	1	30	Number of languages in conflict area, 1945 to 2005.
Cold war	286	0.47	0.00	0.33	0	1	Dummy variable for Cold War, 1949 to 1989.
Duration	283	1.44	7.12	6.89	2.40	9.90	Natural log of days from beginning to end of conflict.

Appendix 2: Replicated Regression Tables

Table 1: *This is the replication of the authors' table 2; foraging and mechanized era of warfare: a comparison. The models were estimated using ordinal logistic regression, with Huber-White robust standard errors clustered on each country in the dataset.*

Variables	Model 1 (MODERN only) (1800-2005)	Model 2 (Full model) (1800-2005)	Model 3 (Foraging) (1800-1917)	Model 4 (Mechanized) (1918-2005)
Modern	-1.90***	-1.77***		
	(0.42)	(0.60)		
Railway			-1.10	
			(0.61)	
Regime		-0.03	-0.02	-0.03
		(0.02)	(0.06)	(0.02)
Support		-0.84***	-1.35***	-0.84***
		(0.16)	(0.34)	(0.17)
Power		0.17	0.57***	0.16
		(0.10)	(0.22)	(0.12)
Energy		0.04	0.13	-0.05
		(0.06)	(0.10)	(0.09)
Occupy		-0.92**	0.04	-1.14***
		(0.36)	(0.74)	(0.34)
Elevation		0.02	-0.37	0.11
		(0.08)	(0.36)	(0.10)
Distance		-0.06	-0.83	-0.03
		(0.06)	(0.43)	(0.06)
Cold War		0.56		0.52
		(0.37)		(0.40)
Cutpoints	-2.37***	-3.35***	-10.65***	-1.05
	(0.27)	(0.72)	(2.61)	(0.86)
	-1.40***	-2.22**	-10.25***	0.36
	(0.24)	(0.68)	(2.59)	(0.90)
N (clusters)	286 (85)	285 (85)	112 (20)	173 (80)
Wald chi^2	20.57***	51.47***	97.27***	32.79***
Log-likelihood	-261.03	-239.56	-54.71	-171.10
R^2	0.09	0.17	0.17	0.10

Robust standard errors clustered on country in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%.

Table 2: *This is the replication of the authors' table 3; the perils of mechanization. The models were estimated using ordinal logistic regression, with Huber-White robust standard errors clustered on each country in the dataset.*

Variables	Model 5 (MECH only) (1918-2005)	Model 6 (MECH) (1918-2005)	Model 7 (MECH) (1945-2005)	Model 8 (HELI) (1945-2005)
Mech	-0.34***	-0.31**	-0.33**	
	(0.13)	(0.14)	(0.16)	
Heli				-1.32**
				(0.44)
Regime		-0.03	-0.05**	-0.05**
		(0.02)	(0.03)	(0.03)
Support		-0.80***	-0.57***	-0.56***
		(0.17)	(0.20)	(0.21)
Power		0.09	0.20	0.38**
		(0.15)	(0.15)	(0.17)
Energy		0.02	0.00	-0.05
		(0.09)	(0.09)	(0.09)
Occupy		-1.32**	-1.76***	-2.20***
		(0.37)	(0.56)	(0.64)
Elevation		0.09	0.18	0.17
		(0.10)	(0.13)	(0.14)
Distance		-0.02	0.03	0.02
		(0.06)	(0.07)	(0.07)
Language			-0.05*	-0.05**
			(0.03)	(0.03)
Trade			0.23	0.28
			(0.34)	(0.30)
Cutpoints	-1.45***	-2.20**	-2.84*	-2.86**
	(0.37)	(0.80)	(1.25)	(1.17)
	-0.22	-0.77	-1.16	-1.14
	(0.35)	(0.78)	(1.21)	(1.13)
N (clusters)	167 (80)	167 (80)	135 (76)	135 (76)
Wald chi^2	7.01***	33.69***	37.49***	42.70***
Log-likelihood	-179.20	-162.96	-129.78	-126.95
R^2	0.02	0.11	0.12	0.14

Robust standard errors clustered on country in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%.