Supplementary Information for
"Inter-Rebel Alliances in the Shadow of Foreign Sponsors"

Contents

List of Tables

List of Figures

Main models in regression table

Table 1 – Pooled Models of Alliance-Making in Civil War, 1975–2009

	Base Model	Model 1	Model 2	Model 3
Single sponsor		0.68	0.34	
		(0.38)	(0.37)	
Different sponsor		1.32	1.09	
		(0.83)	(0.80)	
Shared sponsor		2.27^*	2.14^*	
		(0.39)	(0.38)	
Weak dyad			−0.10	
			(0.82)	
Sponsor			1.46^*	
			(0.32)	
Sponsor x Weak dyad			−0.87	
			(1.21)	
Ratio				2.73^*
				(0.71)
GDP p.c. (ln)	0.17	0.14	0.16^*	0.19^*
	$(0.07)^*$	(0.08)	(0.08)	(0.08)
Expenditure (ln)	0.01	0.02	0.03	0.03
	(0.07)	(0.05)	(0.05)	(0.05)
Duration (ln)	-1.15^*	-1.27^*	-1.24^*	-1.17^*
	(0.15)	(0.15)	(0.15)	(0.15)
Non-Contiguity	−0.08	0.20	0.08	−0.11
	(1.16)	(1.10)	(1.17)	(1.21)
Religious frac.	1.35	0.66	1.31	1.19
	(2.35)	(2.21)	(2.36)	(2.43)
Ethnic frac.	−2.66	−2.44	−2.65	−2.60
	(1.98)	(1.88)	(2.01)	(2.08)
Durability (ln)	0.63^*	0.65^*	0.68^*	0.82^*
	(0.18)	(0.19)	(0.19)	(0.20)
Intercept	−0.77	−1.35	−1.59	−2.78
	(1.4)	(1.37)	(1.42)	(1.56)
Log Likelihood	-289.69	-272.44	-278.18	-280.85
AIC	597.40	568.91	580.39	581.76
BIC	641.41	627.60	639.07	630.64
Num. obs.	985	985	985	985
Num. groups: cowcode	35	35	35	35

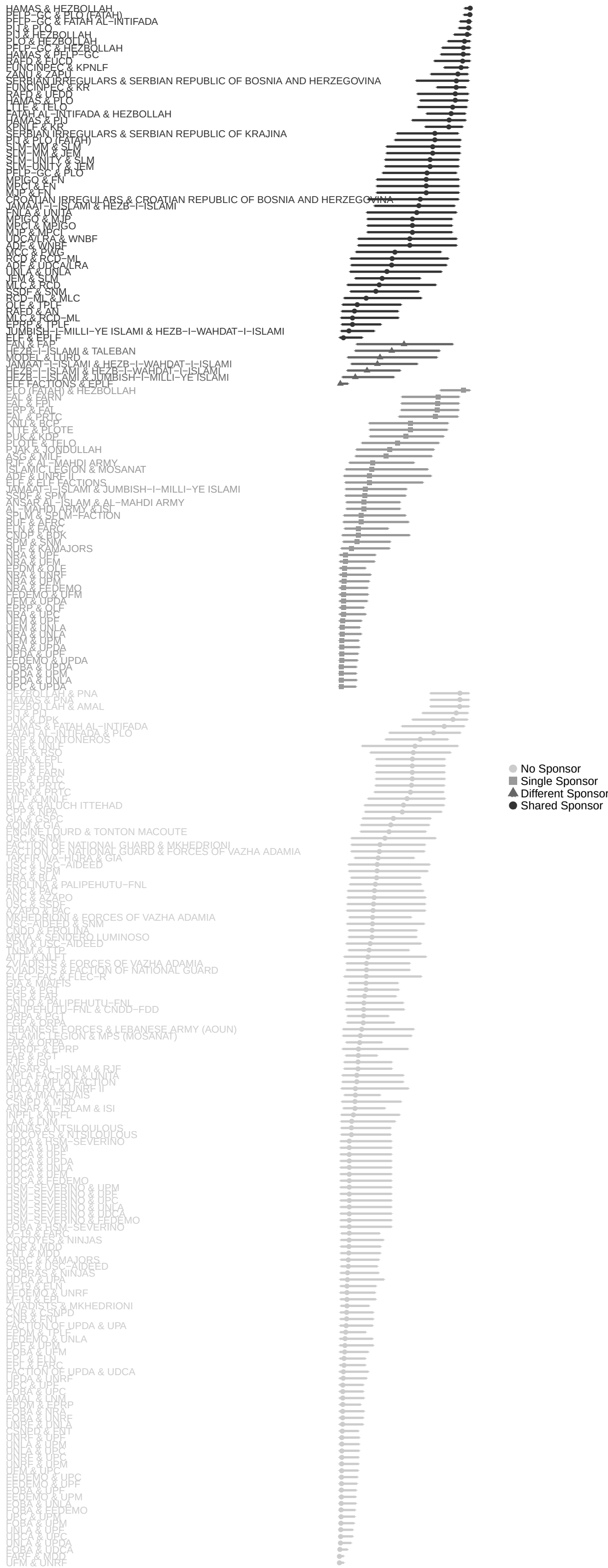
Note: Pooled coefficients estimates from multiple imputations (n=500) for multilevel logistic regression with standard errors in brackets clustered on conflicts. $^*p < 0.05$.

Predicted Probabilities

Figure 1 indicates a high probability of inter-rebel cooperation for the majority of dyads with shared sponsors. My model gives a predicted probability of insurgent alliance of more than 60 percent for roughly two-thirds of dyads in which both partners received external assistance. Notable cases include the FUNCINPEC-KPLNF alliance that was sponsored by the United States, China and Thailand during Cambodian civil war (1980–1991); the ZANU-ZAPU cooperation against Rhodesia with the support of the Eastern bloc and frontline countries (Mozambique and Zambia); and a number of Palestinian groups that received support from Arab countries. Additional cases with high probability of alliance formation include UNITA and FNLA with the support from Mobutu's Zaire; Hezbollah and Hamas who received support from Iran and Syria; major Darfur groups, JEM and SLM, with the backing from Chad; as well as the cooperation between the Tamil Tigers and TELO in the shadow of Indian support in the 1908s.

Roughly a quarter of cases without foreign support have 0.6 or higher probability of alliance. Most of these cases include dyads with much weaker capabilities relative to the incumbent government. This implies that foreign support may not be a necessary catalyst for cooperation where power asymmetry is not a prior issue. The probability of alliance for dyads with single and different sponsors is much lower. While there are only five cases of dyads with single sponsors with probability higher than 0.5, there are no dyads with different sponsors that score a probability of alliance higher than 0.5. This suggests that foreign support seems to be the strongest catalyst for rebel cooperation when there is a shared sponsor.

Figure 1 – Mean Pre 0 percent confidence intervals.



Alternative Controls

In the following models, I include alternative, dyad-level, variables since the main model include only state-level predictors. First, I include shared ideology, which is measured using Non-State Armed Groups (NAG) dataset, I check whether the groups in those dyads shared one of the possible ideologies: left-wing, nationalist, religious, right-wing. If both groups in a dyad followed the same ideology, then "Shared ideology" was coded 1, and 0 otherwise. Next, I code whether any of the groups in a dyad has splintered from the other in the past. To measure "Splinter" I use UCDP Actor Dataset (variable "splinter") which identifies the splintering and parent organization. "Splinter" is coded 1 if one of the parties has splintered from the other in the past, and 0 otherwise. To account for the possibility that shared ethnic/sectarian background drives alliance dynamics, I include "Shared ethnic pool" in the equation. This predictor comes from "recruitment" variable from ACD2EPR dataset Version 2014. This variable indicates whether a rebel group is recruiting from an ethnic group. Possible values are 0 ("no recruitment"), 1 ("recruitment"), 2 ("ethnic group members are recruited by the rebels and the government"). I merge values of 1 and 2 into a single value indicating the presence of ethnic pool (1), or it absence (0). Then I compare if both groups recruit from the same ethnic group. Sometimes rebel groups recruit from numerous ethnic groups. If they recruited from at least one common ethnic group, I coded "Shared ethnic pool" as present or 1, and 0 otherwise. Finally, I include logged duration to account for the temporal dimension of a dyadic relationship. Model 4 features only alternative controls, while Model 5 includes alternative controls and foreign sponsorship.

Table 2 – Alternative controls and rebel alliance (1975–2009)

	Model 4	Model 5
Shared ideology	1.14^{*}	1.15^{*}
	(0.37)	(0.36)
Splinter	0.46	0.32
	(0.45)	(0.47)
Shared ethnic pool	-0.18	-0.24
	(0.45)	(0.45)
Duration (ln)	-1.09^{*}	-1.22^{*}
	(0.15)	(0.15)
Single sponsor		0.34
		(0.37)
Different sponsor		1.09
		(0.80)
Shared sponsor		2.14^{*}
		(0.38)
(Intercept)	-0.53	-1.00^{*}
	(0.48)	(0.46)
Log Likelihood	-298.04	-281.25
AIC	608.08	580.50
BIC	637.44	624.54
Num. obs.	985	985
Num. groups: conflict	35	35

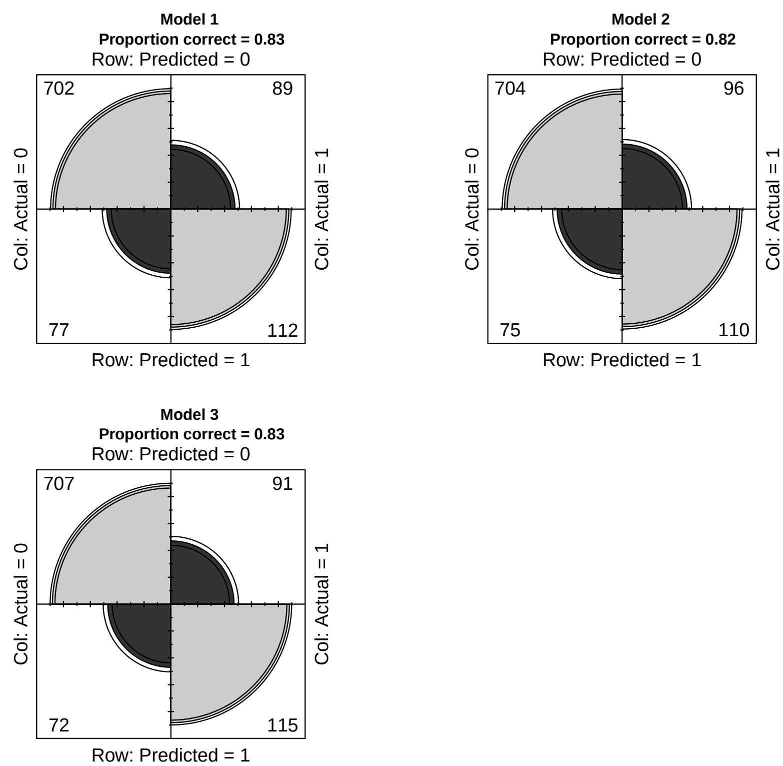
Coefficients estimates for multilevel logistic regression with standard errors in brackets clustered on conflicts. $^{*}p < 0.05$

10-Fold Cross Validation

K-fold cross validation (CV) is a way to analyze how the results of a model apply to an independent sample, i.e. predictive accuracy of the model. The first round of CV includes partitioning of original data into similar folds of subsets (in my case into 10 folds of 98 or 99 observations), performing analysis on a single subset ("training dataset"), and validating the analysis on the other ("testing dataset"). After the data-partitioning, I carry out 500 multiple imputations and then run multilevel logistic model on the training dataset. Next, I loop over the 500 models to obtain predicted values for each fold.

One way of presenting CV results is to make a contingency table composed of two actual classes ("alliance" and "non-alliance") and two predicted classes ("alliance" and "non-alliance") for a set of test data for which true values are known. The northwest (0/0) and southeast (1/1) quadrants are correct predictions, and the other diagonal (northeast and southwest) represent incorrect predictions. The predictive accuracy for each model is .72 (see Figure 2).

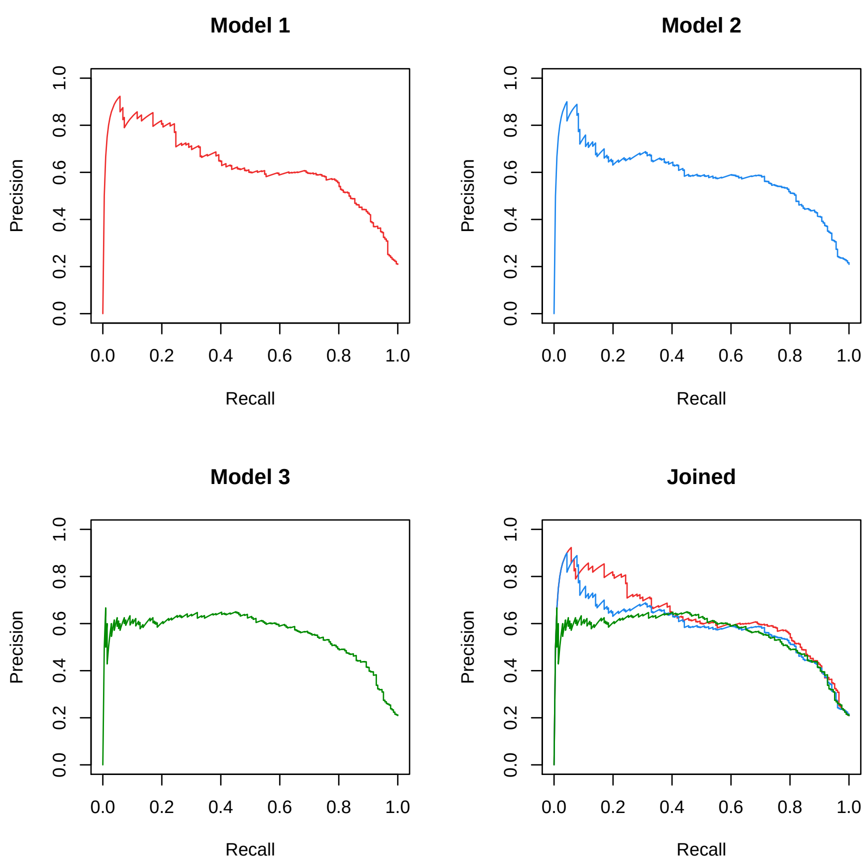
Figure 2 – Predictive Accuracy of Models 1-3 using multilevel logistic regression



The predictions are classified by true negatives in the upper-left quadrant (correct predictions that there is no intervention); true positives in the lower-right quadrant (correct predictions that there is intervention); false positives in the upper-right quadrant (incorrect predictions that there is intervention); and false negatives in the lower-left quadrant (incorrect predictions that there is no intervention). The higher the number of predictions among true/false positives compared to false positives/negatives, the higher the accuracy of our models. Accuracy is the sum of true/false positives divided by the total number of predictions, and shows how often the classifier is correct.

Another way to analyze the predictive accuracy is to make Precision-Recall (PR) curve plots. PR graphs show how meaningful is a positive result ("alliance" occurring) given the baseline probabilities of alliance. PR plots display pairs of recall and precision values where "recall" is a performance measure of the whole positive part of my data (x-axis), whereas "precision" is a performance measure of positive predictions (y-axis). Models with best performance scores are located closer to the the upper-right quadrant. Figure 3 shows that Model 1 has the best performance.

Figure 3 – ROC Curves Predicting Inter-Rebel Alliance



Model Criticism Plot

Figure 4 – Model Criticism Plot for Model 2 (Bapat and Bond)

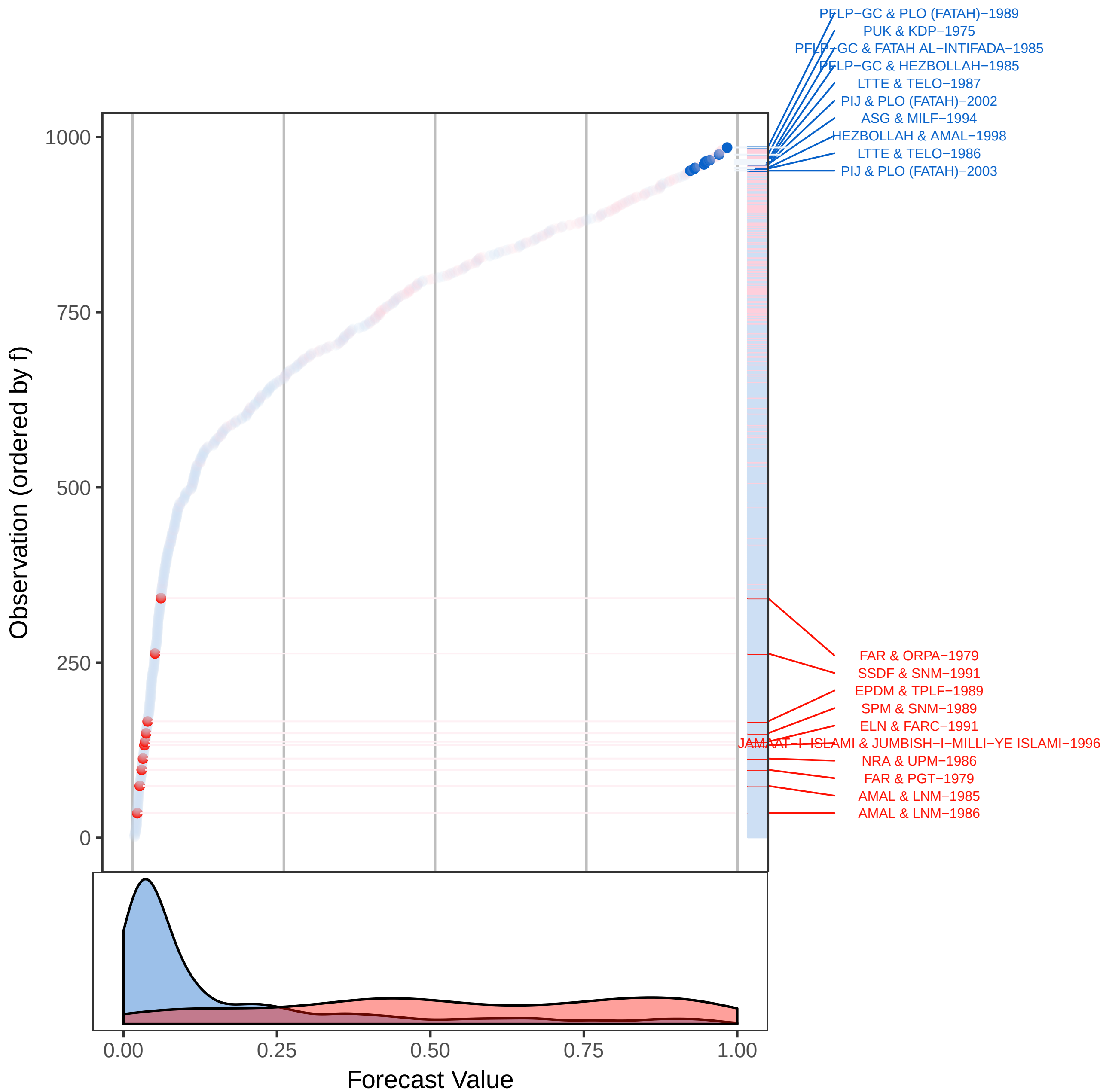
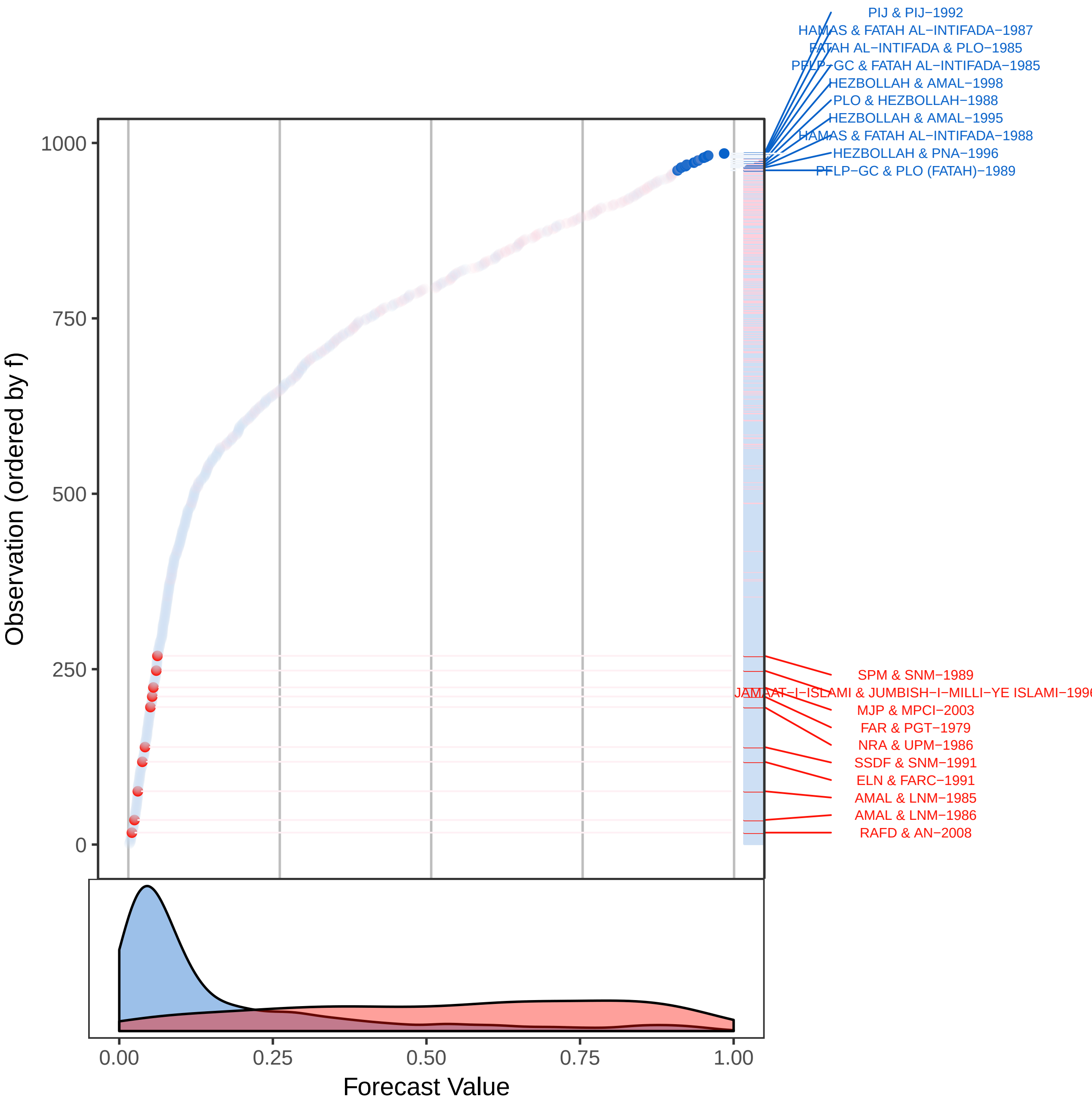


Figure 5 – Model Criticism Plot for Model 3 (Christia)



Selection Bias

Selection bias occurs due to the systematic exclusion of a subset of data due to a specific attribute. The exclusion of the subset will yield distorted empirical results of the population of interest. In my case, it might be that rebel groups that do not expect to enter alliance may have less need for external support, and thus may not appear in my dataset. I cannot analyze alliance-making between such groups because they are systematically omitted from the sample. The issue in my case is that the missingness of foreign support could be a function of rebel groups anticipating to form an alliance. This means that the selection mechanism may be associated with the presence of alliance, yielding biased estimates or estimates that apply only to the selected sub-sample.

Table 3 – Two-Stage Heckman Tobit Model

	Model 6
SELECTION (DV: Foreign support):	
GDP p.c. (ln)	0.13*
	(0.01)
Expenditure (ln)	−0.01
	(0.01)
Weak link	0.04
	(0.37)
Religious frac.	1.71*
	(0.34)
Ethnic frac.	0.90*
	(0.40)
Intercept	−1.99*
	(0.37)
OUTCOME (DV: Alliance):	
Durability (ln)	0.04
	(0.02)
Duration (ln)	−0.19*
	(0.03)
Shared ideology	−0.11
	(0.07)
Shared ethnic ties	0.23*
	(0.07)
Splinter	0.10
	(0.11)
Intercept	0.50*
	(0.10)
Inverse Mills Ratio	−0.14
	(0.09)
sigma	0.42
	NA
rho	−0.34
	NA
R^2	0.26
Adj. R^2	0.23
Num. obs.	536
Censored	358
Observed	178

$^*p < 0.05$

To account for this problem, I use a tobit two-step Heckman selection model (Heckman 1979). I first model the probability of dyads receiving foreign support: in this step, dyads are coded 1 if they received external support from one or more foreign governments (based on UCDP External Support Dataset) and 0 otherwise. The selection model includes a number of variables from Salehyan, Gleditsch and Cunningham (2011) such as logged GDP per capita, strength of the dyad relative to the government, military expenditure and ethnic and religious fractionalization. I then use the information from the first step to compute the Mills inverse ratio and include it in the outcome step. The significance of the coefficient of the Inverse Mills ratio will indicate if there is selection bias. The outcome step includes durability of the countryâĂŹs regime, the logged duration of dyad, shared ideological and ethnic ties, splintered group and Inverse Mills ratio (IMR) as covariates. Table 3 in the Appendix presents the results of the Heckman model. In short, the results show that the IMR is small (beta -0.15) and the p-value is large (0.16), so I cannot reject the null that the errors are uncorrelated.

Robustness Checks

Figure 6 – Pooled Models of Inter-Rebel Alliance, 1975–2001

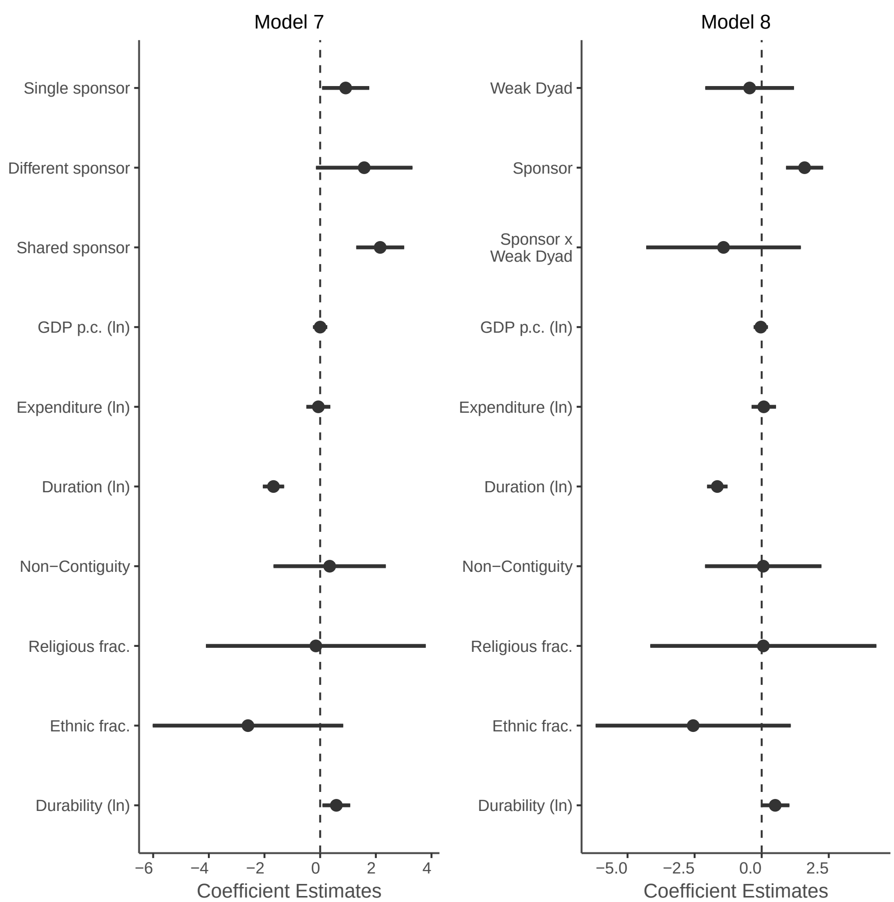


Figure 7 – Bapat and Bond Pooled Model of Inter-Rebel Alliance, 1975–2001, using Bapat and Bond Measure of Foreign Support

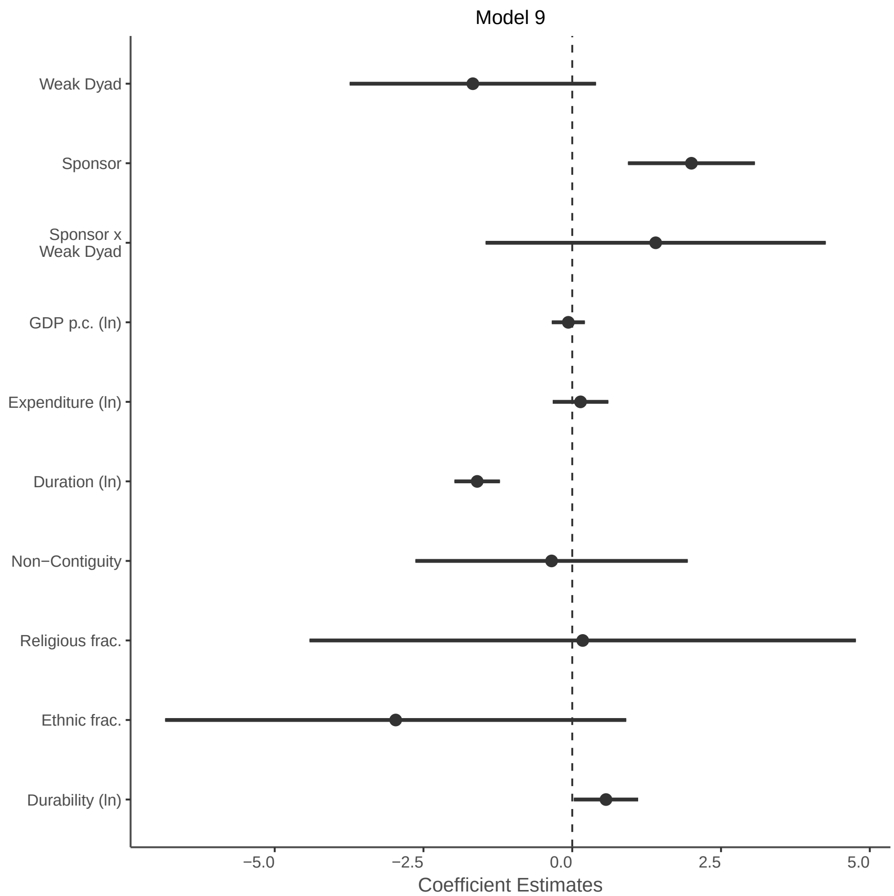


Figure 8 – Rebel-Level Variables and Cold War