Appendix:
Regime Type, Coalition Size, and Victory

Benjamin A. T. Graham* Erik Gartzke† Christopher J. Fariss‡§

Contents

10 Introduction to the Appendix **2**
10.1 Testing Hypotheses 1-3 with Logged Partners 3
10.2 Bivariate Ordered Probit Models 7
10.3 Additional Discussion: Quantity versus Quality in Democratic Coalitions 12
10.4 Coalition Partnership Statistics and Examples 13
10.4.1 Probability of Coalition Partnership 13
10.4.2 Coalition Partnership Example: Korean War (1950-1953) 15
10.4.3 Coalition Partnership Example: Gulf War (1991) 16
10.5 Logging CINC Scores 18
10.6 Temporal Sensitivity Analysis and System-Level Democracy 22
10.7 Controlling for Initiation and its Interaction with Regime Type 28
10.8 Testing Robustness to Reiter, Stam, and Horowitz's Revised War Data 35
10.9 Permutation Analysis for Ordered Probit Models 42
10.9.1 Permutation of Number of Partners Variable 43
10.9.2 Permutation of Logged Number of Partners Variable 44
10.9.3 Permutation of Partners' Summed CINC Scores Variable 45
10.10 Cross Validation for Ordered Probit Models 46
10.10.1 Cross Validation of Number of Partners Variable 48
10.10.2 Cross Validation of Logged Number of Partners Variable 49
10.10.3 Cross Validation of Partners' Summed CINC Scores Variable 50
10.11 Predicted Number of Coalition by CINC Quantile Values 51
10.12 Omitting the Control for Dyad MID Propensity 53
10.13 Splitting the Sample Between Joiners and Initial Participants 56

*University of Southern California.
†University of California, San Diego.
‡Pennsylvania State University
§We would like to thank the participants of the 2011 meeting of the International Studies Association and the 2010 meeting of the International Studies Association West.

10 Introduction to the Appendix

The supplementary material presented in this document provides additional details about the analyses presented in the paper "Regime Type, Coalition Size, and Victory". The main document makes reference to the materials contained here. Replication materials are available here [URL PENDING]. Results from the additional model specifications, coupled with the cross validation tests and permutation tests for the coalition variables presented below, provide substantial evidence that our coalition variables are important predictors of the probability of victory in the international system.

10.1 Testing Hypotheses 1-3 with Logged Partners

In the body of the paper we measure coalition size as the number of partners on side 1 (Hypotheses 1-3) and the aggregate capabilities of those partners (Hypotheses 1a-3a). In this section, we retest Hypotheses 1-3 measuring coalition size as the natural log of the number of states on side 1 (i.e. the log of the number of partners plus one). The results presented in Tables 1-3 are consistent with those we present in the body of the paper.

The only difference between these results and those in the body of the paper is that the effect of coalition size on victory in Table 2 is only significant at p<0.1 in the "Wars Only" sample. However, in the bivariate probit models (Table 3), the same effect is statistically significant at the .05 level across all four models. Taken together, this third set of results is very similar to those the body of the paper, and further girds our confidence that our findings are not driven by the manner in which we measure our variables of interest.

Note that, in Table 1, Models 1 and 2, the dependent variable is not sufficiently over-dispersed to justify negative binomial estimation and the model reduces to a poisson regression.

Table 1: Regime Type and Number of Coalition Partners (logged)

	Wars Only	Wars Only	MIDs (3-5)	MIDs (3-5)
Democracy [Polity IV]	0.029***		0.035***	
	(0.0079)		(0.010)	
Democracy [Boix et al.]		0.53***		0.57***
		(0.11)		(0.15)
CINC score	0.51	0.37	1.06	1.04
	(0.68)	(0.64)	(0.82)	(0.80)
Opponent(s)' CINC score	0.69	0.77*	2.00***	2.03***
	(0.44)	(0.45)	(0.75)	(0.78)
Constant	0.17	-0.021	-1.20***	-1.41***
	(0.14)	(0.14)	(0.16)	(0.12)
ln(alpha)	NA	NA	0.33	0.34*
			(0.22)	(0.21)
Observations	397	410	5019	5229

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$
Specification: Poission Regression (Models 1&2), Negative Binomial regression (Models 3&4).
Errors clustered on dispute-side.

Table 2: Probability of Victory

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
Number of Partners (logged)	0.24*	0.22*	0.35***	0.35***
	(0.13)	(0.13)	(0.089)	(0.088)
Democracy [Polity IV]	0.021		0.012***	
	(0.015)		(0.0038)	
Democracy [Boix et al.]		0.50**		0.21***
		(0.24)		(0.058)
CINC score	5.17***	4.87***	1.98***	2.03***
	(1.65)	(1.55)	(0.47)	(0.46)
Opponent(s)' CINC score	-3.03***	-2.94***	-2.55***	-2.55***
	(0.85)	(0.81)	(0.57)	(0.56)
Dyad MID Propensity	-0.88	-0.86	-0.97	-0.70
	(8.17)	(8.17)	(1.73)	(1.73)
Troop Quality	0.067	0.043	0.020	0.014
	(0.084)	(0.083)	(0.015)	(0.011)
Cut1	-0.22	-0.13	-1.21***	-1.13***
	(0.32)	(0.35)	(0.070)	(0.075)
Cut2	0.50	0.59*	1.24***	1.31***
	(0.31)	(0.32)	(0.062)	(0.064)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Ordered Probit with errors clustered on dispute-side.

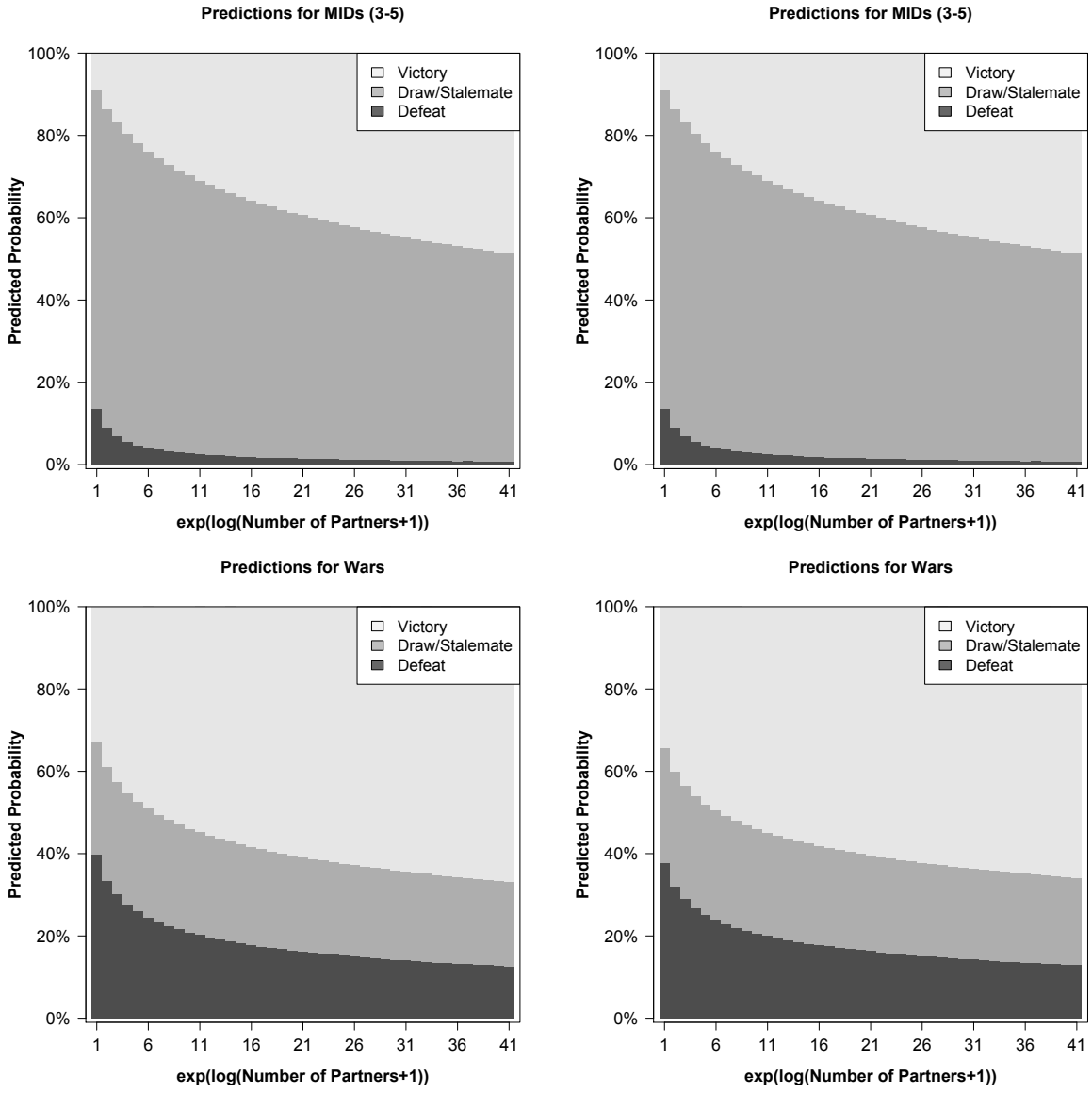

Figure 1: Predicted probability of ordered outcomes during MIDs (3-5) and Wars only as a function of the number of coalition partners using two different measures of democracy (Polity IV in the two left panels and Boix, Miller and Rosato (2013) in the two right panels). All other variables from the models are held at their mean or median values.

Table 3: Joint Probability of Partners and Victory

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = ALLY				
Democracy [Polity IV]	0.047***		0.030***	
	(0.017)		(0.0075)	
Democracy [Boix et al.]		0.69***		0.45***
		(0.25)		(0.11)
CINC score	-0.45	-0.53	-0.36	-0.28
	(1.50)	(1.45)	(0.63)	(0.61)
Opponent(s)' CINC score	1.57	1.51	2.00***	2.02***
	(1.01)	(1.02)	(0.69)	(0.68)
Dyad MID Propensity	14.4	14.9	2.98	3.52
	(9.22)	(9.22)	(3.25)	(3.25)
Constant	-0.44	-0.68*	-1.38***	-1.57***
	(0.34)	(0.35)	(0.13)	(0.12)
DV = WIN				
Number of Partners (logged)	0.58**	0.52**	0.74***	0.73***
	(0.24)	(0.26)	(0.13)	(0.13)
Democracy [Polity IV]	0.000087		-0.0013	
	(0.021)		(0.0054)	
Democracy [Boix et al.]		0.098		-0.081
		(0.34)		(0.082)
CINC score	5.06***	4.79***	2.75***	2.88***
	(1.24)	(1.17)	(0.44)	(0.42)
Opponent(s)' CINC score	-2.84***	-2.72***	-0.50	-0.54
	(0.95)	(0.95)	(0.60)	(0.59)
Troop Quality	0.074	0.070	0.029*	0.024*
	(0.087)	(0.085)	(0.016)	(0.013)
Dyad MID Propensity	-6.06	-4.65	2.72	2.79
	(8.79)	(9.08)	(2.27)	(2.23)
Constant	-0.85***	-0.83**	-1.78***	-1.74***
	(0.32)	(0.35)	(0.075)	(0.078)
Arc-Hyperbolic Tangent	-0.68	-0.58	-0.22	-0.20
	(0.46)	(0.47)	(0.15)	(0.15)
Observations	344	353	4534	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.2 Bivariate Ordered Probit Models

Just as we test the robustness of our findings to different measures of the variables of interest, we also want to test the robustness of the results to different estimation techniques. In this section we model the joint ordered outcomes with correlated errors using a bivariate ordered probit model, implemented using Stata 13.[1] The dependent variable for the two equations for this regression are similar to those in the bivariate probit model discussed in the paper. We use the same dependent variable, *victory* used in tests of Hypothesis 2 and 2a in the body of the paper. In converting coalition size measures to ordinal form, we first create *ordered partners,* which takes a value of 0 if there are no partners, a value of 1 if there are 1 or two partners, and a value of three if there are more than three partners. We also employ *ordered partner capabilities,* which takes a value of zero if there are no partners, 1 if the summed CINC scores of the partners is between 0 and 0.1, and 2 if the summed CINC scores of the partners exceed 0.1.

The same controls used above enter the two equations of this model and errors are clustered on dispute-side. The results corroborate the findings presented in the paper and lend additional support for both links in our argument. They provide direct support for Hypotheses 3 and 3a, respectively. Democracies tend to win the wars that they fight, and they do so predominantly because they fight as part of larger coalitions.

These results are also robust to extending the number of ordered categories regarding the number of allies – we have run bivariate ordered probit models with both four-category and five-category measures and find consistent results.

Tables 4-6 contain a total of 12 bivariate ordered probit models: two different samples, two different measures of democracy, three different measures of coalition size. In all twelve regressions, the effect of democracy on coalition size is positive and statistically significant ($p<.05$). The estimated effect of coalition size on victory is positive in all models, significant at the .05 level in nine, at the 0.1 level in one, and not significant in two.

Model 2 in Table 6 offers the only specification which runs at least partially counter to

[1] See Calhoun (1986, 1989) and Greene and Hensher (2010) for more details about this model.

our theoretical assertions: in this model alone (out of 24 bivariate models presented in the paper and the appendix and several other robustness tests referenced in these sections but not presented) we see a statistically significant effect of regime type on the probability of victory that is statistically stronger than the effect of coalition size. Therefore, while we do not want to sweep this result under the rug, we feel it does not challenge our overall assertion that the effect of regime type on the probability of victory runs primarily through the tendency of democracies to fight alongside larger and more capable sets of coalition partners.

Table 4: Joint Probability of Partners and Victory: Bivariate Ordered Probit

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Ordered Partners				
Democracy [Polity IV]	0.035**		0.018***	
	(0.016)		(0.0061)	
CINC score	-0.68	-0.53	1.80***	1.83***
	(1.35)	(1.29)	(0.48)	(0.47)
Opponent(s)' CINC score	1.55	1.51	1.87***	1.88***
	(1.01)	(1.01)	(0.54)	(0.54)
Dyad MID Propensity	8.51	8.51	2.22	2.40
	(7.76)	(7.76)	(2.41)	(2.39)
Democracy [Boix et al.]		0.52**		0.27***
		(0.25)		(0.091)
Cut 1	-0.29	-0.084	0.84***	0.94***
	(0.27)	(0.26)	(0.096)	(0.086)
Cut 2	0.26	0.44*	1.40***	1.50***
	(0.26)	(0.26)	(0.12)	(0.10)
DV = Victory				
Number of Partners	0.043**	0.038*	0.067***	0.068***
	(0.021)	(0.020)	(0.014)	(0.014)
Democracy [Polity IV]	0.020		0.010***	
	(0.017)		(0.0037)	
CINC score	5.21***	4.90***	2.34***	2.38***
	(1.65)	(1.55)	(0.46)	(0.45)
Opponent(s)' CINC score	-2.98***	-2.90***	-2.44***	-2.44***
	(0.85)	(0.83)	(0.55)	(0.55)
Troop Quality	0.055	0.033	0.023	0.016
	(0.079)	(0.078)	(0.014)	(0.011)
Dyad MID Propensity	-2.19	-1.79	-1.14	-0.88
	(8.03)	(8.05)	(1.74)	(1.74)
Democracy [Boix et al.]		0.47*		0.18***
		(0.26)		(0.057)
Cut 1	-0.34	-0.25	-1.23***	-1.17***
	(0.33)	(0.35)	(0.071)	(0.075)
Cut 2	0.37	0.47	1.22***	1.28***
	(0.35)	(0.34)	(0.063)	(0.065)
Arc-Hyperbolic Tangent	-0.024	-0.0040	0.0031	0.0029
	(0.19)	(0.19)	(0.066)	(0.066)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Joint Probability of Partners and Victory: Bivariate Ordered Probit

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Ordered Partners				
Democracy [Polity IV]	0.034**		0.018***	
	(0.016)		(0.0061)	
CINC score	-0.49	-0.38	1.79***	1.83***
	(1.40)	(1.33)	(0.49)	(0.47)
Opponent(s)' CINC score	1.56	1.52	1.84***	1.85***
	(1.04)	(1.03)	(0.56)	(0.56)
Dyad MID Propensity	8.27	8.20	2.02	2.21
	(7.74)	(7.73)	(2.41)	(2.39)
Democracy [Boix et al.]		0.49*		0.27***
		(0.26)		(0.091)
Cut 1	-0.30	-0.100	0.82***	0.93***
	(0.27)	(0.26)	(0.098)	(0.088)
Cut 1	0.25	0.42*	1.39***	1.50***
	(0.26)	(0.25)	(0.12)	(0.11) heightDV = Victory
Number of Partners (logged)	0.55**	0.49*	0.55***	0.56***
	(0.25)	(0.27)	(0.14)	(0.14)
Democracy [Polity IV]	0.0075		0.0085**	
	(0.018)		(0.0038)	
CINC score	5.03***	4.77***	1.83***	1.87***
	(1.45)	(1.41)	(0.47)	(0.46)
Opponent(s)' CINC score	-3.19***	-3.13***	-2.73***	-2.74***
	(0.85)	(0.84)	(0.58)	(0.58)
Troop Quality	0.050	0.032	0.021	0.015
	(0.078)	(0.078)	(0.014)	(0.011)
Dyad MID Propensity	-4.30	-3.72	-1.42	-1.21
	(8.07)	(8.11)	(1.72)	(1.70)
Democracy [Boix et al.]		0.29		0.15**
		(0.27)		(0.058)
Cut 1	0.11	0.087	-1.14***	-1.08***
	(0.43)	(0.43)	(0.084)	(0.086)
Cut 2	0.78**	0.78**	1.27***	1.32***
	(0.34)	(0.34)	(0.064)	(0.065)
Arc-Hyperbolic Tangent	-0.47	-0.40	-0.23**	-0.24**
	(0.38)	(0.37)	(0.11)	(0.11)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Joint Probability of Partners and Victory: Bivariate Ordered Probit

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Ordered Partner Capabilities				
Democracy [Polity IV]	0.058***		0.021***	
	(0.015)		(0.0065)	
CINC score	-0.43	-0.34	1.60***	1.67***
	(1.21)	(1.15)	(0.53)	(0.52)
Opponent(s)' CINC score	1.43**	1.36**	1.67***	1.66***
	(0.63)	(0.64)	(0.46)	(0.46)
Dyad MID Propensity	-0.51	0.48	0.85	1.27
	(7.11)	(7.13)	(2.48)	(2.47)
Democracy [Boix et al.]		0.87***		0.29***
		(0.21)		(0.098)
Cut 1	-0.25	0.067	0.97***	1.09***
	(0.23)	(0.23)	(0.095)	(0.085)
Cut 2	0.77***	1.07***	1.44***	1.56***
	(0.19)	(0.18)	(0.11)	(0.090)
DV = Victory				
Partner(s)' CINC score	1.18	0.66	2.38***	2.35***
	(1.07)	(0.98)	(0.81)	(0.80)
Democracy [Polity IV]	0.019		0.013***	
	(0.015)		(0.0040)	
Democracy [Boix et al.]		0.53**		0.22***
		(0.24)		(0.062)
CINC score	5.23***	4.84***	1.86***	1.93***
	(1.73)	(1.61)	(0.49)	(0.48)
Opponent(s)' CINC score	-3.01***	-2.81***	-2.37***	-2.37***
	(0.86)	(0.82)	(0.55)	(0.54)
Troop Quality	0.076	0.050	0.030**	0.021**
	(0.082)	(0.080)	(0.013)	(0.010)
Dyad MID Propensity	1.00	0.94	-0.30	-0.053
	(7.96)	(8.01)	(1.74)	(1.73)
Cut 1	-0.38	-0.29	-1.25***	-1.17***
	(0.34)	(0.34)	(0.075)	(0.075)
Cut 2	0.34	0.43	1.19***	1.26***
	(0.34)	(0.34)	(0.079)	(0.073)
Arc-Hyperbolic Tangent	0.22	0.25	0.10	0.10
	(0.18)	(0.19)	(0.072)	(0.072)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.3 Additional Discussion: Quantity versus Quality in Democratic Coalitions

Here we provide a more detailed discussion about our argument about quantity compared to a related argument about quality. To win on the battlefield because of their coalition partners, democracies must have more coalition partners (our argument), better quality partners (Choi's argument), or some combination of the two. Even if democracies are more effective coalition partners, the democratic reliability argument might still suffer from an individual fallacy; the source of democratic victory could be the size of their coalition, *even if* democracies happen to be better partners.[2] Choi (2004) recognizes the need to address the direct effects of coalition size on war-fighting, but while her research design assumes that reliability (measured as regime type) is uncorrelated with the size of coalitions, her theory must assume in effect that this is not the case. If democracies are better partners, then democratic coalition ties are likely to differ from those of their autocratic counterparts. Democracies should be more sought after because they make better coalition partners. At the same time, democracies have less need to form large coalitions to protect themselves, given the reliability of democratic partners. It is essential to determine the net effect of these contrasting supply and demand effects in order to assess whether it is quantity or quality driving democratic battlefield performance. This is extremely difficult to accomplish theoretically (Conybeare, 1994). Fortunately, however, we can eliminate one set of circumstances in light of available evidence.

High democratic reliability could substitute for coalition size; if partners are more reliable, a state needs fewer of them. Were democracies to tend to partner together, this would lead to the expectation that democracies don't need large coalitions. However, this is not consistent with our finding that democracies have larger wartime coalitions. It is also inconsistent with Choi (2004, Table 1, page 671 and 676), who finds no significant relationship for an interaction term between the regime type of a state at war and the number of democracies in its coalition. This indicates that democracies do not win because they are more likely than autocracies to

[2] Autocracies of equal capability are less effective at aiding an ally (Choi 2004, Figure 1), but the fact that autocracies have a positive marginal impact implies a tradeoff between quantity and quality.

have democracies as their coalition partners. Thus, in addition to ruling out the possibility that partner quality is a substitute for quantity, these findings significantly cloud the first logical connection between democratic coalition efficacy and battlefield victory required for the partner quality argument to function; democracies are not showing a clear preference for other democracies in forming coalitions. If the effect of partnerships with democracies on battlefield victory is not uniquely tied to democratic states, then the tendency of democracies to win contests cannot be attributed to the unique interactions within democracy-dominated coalition structures.

The remaining pathway for the "quality" argument is to assert the dominance of the demand-side. If democratic partners are more desirable than autocratic partners, then democracies should be sought out more often as partners and therefore enter into conflicts in larger coalitions. This implies that, in a given dyad, the likelihood of a coalition being formed is increased when either or both states are democratic. As we have already noted, the literature in this area suggest that democracies do not form significantly more partnerships than non-democracies. Given the demand-side version of the reliability argument, however, we would expect to see that the lowest propensity to partner would be in autocratic-autocratic dyads, and that an intermediate propensity to partner would occur in mixed democratic-autocratic dyads. We assess each of these relationships below.

10.4 Coalition Partnership Statistics and Examples

10.4.1 Probability of Coalition Partnership

As we discussed above and in the literature review of the manuscript, current thinking is that democracies are likely to co-ally, but that autocracies show a similar preference for co-alliance, suggesting that democracies are not universally preferred as allies. We demonstrate this point empirically here. To conduct this supplementary analysis, we are forced to deviate from our focus on *de facto* wartime coalitions and look at *de jure* alliance ties, which can be observed outside of wartime.

Table 7 displays the probability that an alliance exists in any given dyad-year. Alliance data

are from Gibler and Sarkees (2004), as first created by Singer and Small (1966). We utilize a dummy variable for the existence of any alliance in the dyad year, but the substantive results are unchanged if we restrict the analysis to alliances with mutual-defense pact guarantees. We use logistic regression to determine the probability that an alliance exists in a dyad conditional on the dyad type in addition to controls for temporal dependence and military capabilities. Confidence intervals are generated from standard errors, clustered by a dyad number based on COW country codes.

Table 7: Probability of Coalition Partnership

	Probability	95% CI
Democracy-Democracy Dyads	9.40%	[9.26%, 9.54%]
Democracy-Autocracy Dyads	3.00%	[2.96%, 3.05%]
Autocracy-Autocracy Dyads	7.73%	[7.65%, 7.82%]

The results in Table 7 are *inconsistent* with a universal preference for democratic, as opposed to autocratic, allies. Instead, we see evidence that regimes ally with like regimes: democracies with democracies and autocracies with autocracies. This result has been demonstrated elsewhere and should not be considered controversial (Siverson and Emmons, 1991; Simon and Gartzke, 1996; Lai and Reiter, 2000).

While it remains possible that democracies are, in fact, superior coalition partners, the results in Table 7 provide evidence that it is not the superior performance of democratic partners that drives the tendency of democracies to go to war in large coalitions. Nor is it the case that the putatively superior effects of democracies in war fighting are uniquely, or even significantly, associated with other democracies. If democracies tend to win contests because they are fighting in coalition with other democracies, then this must occur through channels that neither favor democratic combatants, nor lead democracies to economize with their coalition ties.

10.4.2 Coalition Partnership Example: Korean War (1950-1953)

In addition to the United States, the following countries contributed forces to the UN Peace-keeping operation in the Korean War (the country's polity score is listed on the right):

Table 8: Coalition Partners in the Korean War (1950-1953)

Country	Polity
Australian	10
Belgium	10
Canada	10
Colombia	1
Denmark	10
Ethiopia	0
France	10
Greece	7
Holland Netherlands	10
India	9
Italy	10
Luxembourg	-
New Zealand	10
Norway	10
Philippines	6
Republic of South Korea	1
South Africa	7
Sweden	10
Thailand	1
Turkey	7
United Kingdom	10

Involvement of these countries was explicitly about opposing aggression. Only India, the Philippines, Thailand and of course ROK are from the region. Out of 22 members of the coalition, 16 are democracies and only 4 are non-democracies, with Luxembourg probably a democracy (not coded) and the Philippines an anocracy. Further, only two autocracies appear from outside the region (Columbia and Ethiopia), while most of the democracies are distant states.

One might argue that most of these countries are not contributing substantial numbers of troops. However, the point then is that the threshold for participation is low and failure to

participate is more illustrative than involvement. If the number of troops needed to contribute is low, one might expect autocracies to participate, especially given that often their costs for fighting are believed to be relatively low. Why then so few autocracies in the Korean War?

Perhaps we can explain this by pointing out that many of the autocracies are socialist countries that naturally do not want to help Western liberal institutions. However, these autocracies are also not fighting on the North Korean side. Russia has an indirect role and the Chinese of course participate in the second half of the war. Beyond this, participation is not widespread. North Korea is fighting to gain territory. South Korea is fighting to prevent this (and sometimes to gain territory themselves). The rest of the participants are fighting for some other objective – they are **not** fighting for a share of the disputed territory. The most attractive reason appears to be defending the principle of territorial integrity (i.e. opposition to aggressive force). Interestingly, late in the war, when South Korea pushed to fight for more territorial concessions, but none of its partners would support this. South Korea may have been motivated by rival goods in this conflict, but its (mostly democratic) coalition partners were not.

10.4.3 Coalition Partnership Example: Gulf War (1991)

As we describe in the main manuscript, for the 31 Gulf War participants with data, the mean Polity score is 6.29, slightly higher than average polity in 1991, but not all that impressive. If, however, we eliminate countries in the immediate region of the conflict that have more parochial reasons to participate in the war, the average Polity score among the 22 out-of-region countries is 9.29 (on a scale from -10 to 10). It should also be emphasized that the mostly democratic distant partners seemed to have no major difficulty working alongside regional autocracies, presumably because their objectives (reinforcing the status quo) were not incompatible with regional actors who might have been concerned with territory.

Table 9: Coalition Partners in the Gulf War (1991)

Country	Polity
Argentina	7
Australia	10
Bahrain	0
Bangladesh	6
Belgium	10
Canada	10
Czechoslovakia	8
Denmark	10
Egypt	0
France	9
Greece	10
Hungary	10
Italy	10
Kuwait	0
Morocco	0
Netherlands	10
New Zealand	10
Niger	-
Norway	10
Oman	0
Pakistan	8
Poland	8
Qatar	0
Saudi Arabia	0
Senegal	2
South Korea	7
Spain	10
Sweden	10
Syria	0
United Arab Emirates	0
United Kingdom	10
United States	10

10.5 Logging CINC Scores

Two of the important control variables in our analysis are the material capabilities of the state in question (*CINC Score* and the capabilities of their opponent(s) (*Opponent(s)' CINC Score*). These variables are right-skewed and it is possible to reduce that skew by taking the natural log of each. In Tables 10-13 we test the robustness of our results to these alternative specifications of the key control variables and find that our results are robust. All the results in Tables 10 and 11 match those in Tables 1 and 2 in the body of the paper, with all variables of interest significant in the expected direction ($p<.05$).

Tables 12 and 13 present the bivariate probit and bivariate ordered probit results, respectively. In the bivariate probit models (Table 12), the effect of regime type on coalition size is positive and statistically significant across all four models, but in the bivariate ordered probit (Table 13) the effect of the binary measure of democracy on coalition size is only significant at the .1 level. Conversely, while the effect of coalition size on the probability of victory is positive and statistically significant ($p < .05$) in all four bivariate ordered probit models (Table 13), that effect falls below that significance threshold in the wars only models in the bivariate probit models (Table 12). This is similar to what occurs in the bivariate ordered probit regressions in Table 5 above, where the CINC variables are not logged.

In Tables 10-13, all effects are in the expected direction and almost all are statistically significant at conventional levels. Taken together, these results further increase our confidence in our results and further demonstrate that the results we observe are not sensitive to the measurement or estimation strategy we adopt.

Table 10: Number of Coalition Partners

	Wars Only	Wars Only	MIDs (3-5)	MIDs (3-5)
Democracy [Polity IV]	0.037***		0.058***	
	(0.011)		(0.016)	
Democracy [Boix et al.]		0.78***		1.03***
		(0.19)		(0.23)
CINC score (logged)	0.034	0.013	-0.0025	-0.0092
	(0.047)	(0.044)	(0.050)	(0.041)
Opponent(s)' CINC score (logged)	0.15*	0.16**	0.11	0.12
	(0.084)	(0.077)	(0.075)	(0.075)
Constant	2.31***	1.97***	0.79*	0.41
	(0.32)	(0.36)	(0.42)	(0.45)
ln(alpha)	0.45*	0.43*	2.11***	2.10***
	(0.23)	(0.23)	(0.10)	(0.099)
Observations	397	410	5019	5229

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$
Specification: Negative Binomial regression with errors clustered on dispute-side.

Table 11: Probability of Victory

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
Number of Partners	0.042**	0.038**	0.064***	0.064***
	(0.017)	(0.016)	(0.014)	(0.014)
Democracy [Polity IV]	0.028**		0.012***	
	(0.014)		(0.0040)	
Democracy [Boix et al.]		0.60**		0.20***
		(0.24)		(0.063)
CINC score (logged)	0.15**	0.13**	0.083***	0.084***
	(0.060)	(0.058)	(0.015)	(0.015)
Opponent(s)' CINC score (logged)	-0.17**	-0.16**	-0.096***	-0.096***
	(0.071)	(0.071)	(0.019)	(0.019)
Dyad MID Propensity	-3.48	-3.01	-1.29	-1.05
	(8.48)	(8.52)	(1.76)	(1.75)
Troop Quality	0.087	0.068	0.027*	0.020*
	(0.076)	(0.075)	(0.014)	(0.011)
Cut 1	-0.30	-0.11	-1.11***	-1.04***
	(0.50)	(0.52)	(0.13)	(0.14)
Cut 2	0.38	0.57	1.32***	1.38***
	(0.53)	(0.53)	(0.14)	(0.14)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Ordered Probit with errors clustered on dispute-side.

Table 12: Joint Probability of Partners and Victory: Bivariate Probit

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = ALLY				
Democracy [Polity IV]	0.044***		0.026***	
	(0.017)		(0.0076)	
Democracy [Boix et al.]		0.63**		0.39***
		(0.27)		(0.11)
CINC score (logged)	-0.082	-0.091	-0.00087	0.00038
	(0.069)	(0.068)	(0.022)	(0.021)
Opponent(s)' CINC score (logged)	0.17**	0.17**	0.11***	0.11***
	(0.082)	(0.083)	(0.038)	(0.038)
Dyad MID Propensity	14.7	15.2	3.07	3.56
	(9.33)	(9.33)	(3.32)	(3.30)
Constant	-0.054	-0.33	-0.79***	-0.94***
	(0.49)	(0.52)	(0.18)	(0.19)
DV = WIN				
Number of Partners	0.045*	0.040	0.091***	0.092***
	(0.027)	(0.027)	(0.018)	(0.018)
Democracy [Polity IV]	0.017		-0.0025	
	(0.022)		(0.0058)	
Democracy [Boix et al.]		0.35		-0.12
		(0.34)		(0.093)
CINC score (logged)	0.17***	0.16***	0.16***	0.16***
	(0.063)	(0.061)	(0.020)	(0.019)
Opponent(s)' CINC score (logged)	-0.12	-0.097	-0.010	-0.011
	(0.074)	(0.075)	(0.021)	(0.021)
Troop Quality	0.12	0.11	0.028*	0.024*
	(0.080)	(0.078)	(0.016)	(0.013)
Dyad MID Propensity	-5.66	-4.67	2.54	2.53
	(9.40)	(9.47)	(2.32)	(2.30)
Constant	-0.15	-0.25	-0.84***	-0.77***
	(0.54)	(0.54)	(0.15)	(0.16)
Arc-Hyperbolic Tangent	-0.20	-0.17	0.076	0.086
	(0.28)	(0.28)	(0.12)	(0.12)
Observations	344	353	4534	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Joint Probability of Partners and Victory: Bivariate Ordered Probit

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Ordered Partners				
Democracy [Polity IV]	0.033**		0.015***	
	(0.016)		(0.0060)	
Democracy [Boix et al.]		0.47*		0.23**
		(0.27)		(0.092)
CINC score (logged)	-0.097	-0.098*	0.065***	0.059***
	(0.059)	(0.058)	(0.018)	(0.018)
Opponent(s)' CINC score (logged)	0.18**	0.18**	0.087***	0.089***
	(0.073)	(0.074)	(0.026)	(0.026)
Dyad MID Propensity	8.42	8.46	1.98	2.13
	(7.37)	(7.43)	(2.41)	(2.39)
Cut 1	-0.64*	-0.44	-0.076	0.036
	(0.38)	(0.40)	(0.13)	(0.14)
Cut 2	-0.075	0.097	0.48***	0.59***
	(0.39)	(0.41)	(0.14)	(0.14)
DV = Victory				
Number of Partners	0.045*	0.040*	0.067***	0.067***
	(0.026)	(0.024)	(0.016)	(0.016)
Democracy [Polity IV]	0.028*		0.012***	
	(0.016)		(0.0039)	
Democracy [Boix et al.]		0.59**		0.20***
		(0.25)		(0.062)
CINC score (logged)	0.15**	0.13**	0.083***	0.084***
	(0.060)	(0.058)	(0.015)	(0.015)
Opponent(s)' CINC score (logged)	-0.17**	-0.16**	-0.096***	-0.097***
	(0.077)	(0.078)	(0.020)	(0.020)
Troop Quality	0.087	0.069	0.028*	0.020*
	(0.077)	(0.076)	(0.014)	(0.011)
Dyad MID Propensity	-3.69	-3.11	-1.32	-1.09
	(8.59)	(8.61)	(1.78)	(1.77)
Cut 1	-0.28	-0.10	-1.10***	-1.04***
	(0.56)	(0.55)	(0.14)	(0.14)
Cut 2	0.40	0.58	1.32***	1.38***
	(0.58)	(0.57)	(0.14)	(0.15)
Arc-Hyperbolic Tangent	-0.036	-0.017	-0.026	-0.023
	(0.21)	(0.21)	(0.070)	(0.069)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.6 Temporal Sensitivity Analysis and System-Level Democracy

It is possible that the structural consequences of the allied victory in WWII or U.S. victory in the Cold War, or the general expansion of democracy over time may have some influence on our results. Therefore, in the following models we control for the proportion of countries in the world that are democratic in any given year (based on the Boix, Miller and Rosato (2013) measure of democracy), and we analyze the interaction between our variables of interest (democracy and coalition size) and dummy variables for the post-Cold War era and the post-WWII era. These interaction terms allow us to assess the degree to which the effects of our variables of interest vary across different time periods.

We create two dummy variables, *Post World War II* and *Post Cold War*, which are dummy variables for post-1945 and post-1989 respectively. Table 14 includes the interaction between both of these dummy variables and democracy. In this table, the main effect of regime type can be interpreted as the effect of democracy on coalition size in the pre-1946 period. Similarly, Table 15 includes the interaction between both of these dummy variables and coalition size. In this table, the main effect of coalition size can be interpreted as the effect of coalition size on victory in the pre-1946 period. The main effects of democracy and coalition size retain the same interpretation in the jointly estimated models (Tables 16 and 17).

In the bivariate probit models (Table 16), we are only able to report results from regressions on the sample of high intensity MIDs. If we restrict the sample to wars only the models won't converge – likely because there are few wars in the post-Cold War sample, making estimation of the interaction term difficult.

The results in Table 14 show positive coefficients, of borderline statistical significance, on the interaction between democracy and the post-Cold War dummy variable, and positive, statistically insignificant coefficients on the interaction between democracy and the post-WWII dummy. This suggest that the effect of democracy on the number of allies is weakest in the period before WWII, but the estimated effect is positive in all periods.

These results are born out in the jointly estimated models (Tables 15 and 16). Consistent with our theory, the estimated effect of democracy on coalition size in the pre-WWII period

remains positive in every model, and is statistically significant in several, including both of the bivariate probit models we consider our "primary" specifications. The relationship between regime type and coalition size is strongest in more recent periods, but positive throughout the full sample we examine.

Our results with regard to the effect of coalition size on victory are even more stable over time. While the average size of coalitions is generally smaller in the post-WWII period, the effect of coalition size on victory varies little from time period to time period. Across all the models estimated, we fail to see a single statistically significant coefficient on any of the interaction terms between coalition size and either the post-Cold War or Post-WWII period. This means that we cannot distinguish between effects of coalition size and the probability of victory between each time period.

These temporal sensitivity results corroborate the average effects reported in all of the models that do not include the temporal dummy variables and interactions. Moreover, these results, coupled with the cross validation tests and permutation tests for the coalition variables presented below, provide even more evidence that the coalition variables are important predictors of the probability of victory in the international system.

Similarly, we draw additional confidence from the robustness of our results to the inclusion of the control for the average level of democracy in the system in any given year. As expected, coalitions are larger in years where there are more democracies in the system, but this system-level effect does not subsume the country-level effect at the core of our results. The estimated effect of regime type on victory remains strong across all these models.

Table 14: Number of Coalition Partners

	Wars Only	Wars Only	MIDs (3-5)	MIDs (3-5)
Democracy [Polity IV]	0.00044		0.0084	
	(0.0080)		(0.0076)	
Democracy [Boix et al.]		0.27*		0.38**
		(0.15)		(0.17)
CINC score	1.16	0.79	0.50	0.23
	(0.91)	(0.90)	(0.75)	(0.82)
Opponent(s)' CINC score	-0.95	-0.85	2.61**	2.58**
	(1.23)	(1.20)	(1.28)	(1.24)
System Level Democratic Share [Boix]	7.55***	7.17***	3.74***	3.53***
	(1.69)	(1.74)	(1.21)	(1.15)
Post World War II	-1.17**	-1.14**	-0.84***	-0.85***
	(0.46)	(0.46)	(0.30)	(0.30)
Democracy [Polity] * Post WWII	0.043**		0.016	
	(0.019)		(0.015)	
Post Cold-War	-1.78***	-1.85***	0.64	0.32
	(0.59)	(0.63)	(0.49)	(0.47)
Democracy [Polity] * Post Cold War	0.044		0.059*	
	(0.034)		(0.033)	
Democracy [Boix et al.] * Post WWII		0.031		0.0063
		(0.019)		(0.014)
Democracy [Boix et al.] * Post Cold War		0.61		0.89**
		(0.41)		(0.36)
Constant	0.049	0.067	-0.95***	-0.99***
	(0.42)	(0.42)	(0.25)	(0.26)
ln(alpha)	0.28	0.26	1.98***	1.98***
	(0.24)	(0.25)	(0.12)	(0.11)
Observations	397	395	5019	4992

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$
Specification: Negative Binomial regression with errors clustered on dispute-side.

Table 15: Probability of Victory

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
Number of Partners	0.071**	0.064**	0.066**	0.068**
	(0.028)	(0.027)	(0.027)	(0.029)
Democracy [Polity IV]	0.023*		0.011***	
	(0.014)		(0.0036)	
Democracy [Boix et al.]		0.49**		0.19***
		(0.21)		(0.056)
CINC score	5.12***	4.77***	2.22***	2.19***
	(1.80)	(1.72)	(0.45)	(0.45)
Opponent(s)' CINC score	-3.33***	-3.34***	-2.46***	-2.50***
	(0.86)	(0.85)	(0.55)	(0.54)
Dyad MID Propensity	-3.01	-2.64	-1.45	-1.06
	(8.05)	(7.93)	(1.69)	(1.66)
Troop Quality	0.065	0.039	0.023*	0.016
	(0.080)	(0.079)	(0.013)	(0.010)
System Level Democratic Share [Boix]	-0.041	0.42	-0.066	-0.027
	(1.99)	(1.96)	(0.54)	(0.53)
Post World War II	0.29	0.099	0.0078	-0.031
	(0.47)	(0.47)	(0.096)	(0.095)
Post Cold-War	-0.15	-0.20	-0.15	-0.14
	(0.51)	(0.50)	(0.11)	(0.11)
Partners * Post WWII	-0.060	-0.056	-0.0044	-0.0073
	(0.037)	(0.037)	(0.035)	(0.036)
Partners * Post Cold War	0.018	0.022	0.014	0.015
	(0.040)	(0.039)	(0.028)	(0.028)
Cut 1	-0.30	-0.17	-1.29***	-1.23***
	(0.61)	(0.61)	(0.18)	(0.18)
Cut 2	0.42	0.56	1.16***	1.22***
	(0.62)	(0.61)	(0.18)	(0.18)
Observations	336	345	4104	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Ordered Probit with errors clustered on dispute-side.

Table 16: Joint Probability of Partners and Victory: Bivariate Probit

	(1) MIDs (3-5)		(2) MIDs (3-5)	
DV = ALLY				
Democracy [Polity IV]	0.017**	(0.0081)		
Democracy [Boix et al.]			0.25**	(0.11)
CINC score	-0.093	(0.48)	-0.21	(0.49)
Opponent(s)' CINC score	1.94***	(0.63)	1.89***	(0.63)
Dyad MID Propensity	3.19	(3.26)	3.42	(3.29)
System Level Democratic Share [Boix]	1.40*	(0.73)	1.42*	(0.75)
Post World War II	-0.49***	(0.17)	-0.53***	(0.18)
Post Cold-War	0.17	(0.23)	0.039	(0.25)
Democracy [Polity] * Post WWII	0.0000093	(0.013)		
Democracy [Polity] * Post Cold War	0.042**	(0.019)		
Democracy [Boix et al.] * Post WWII			0.0072	(0.011)
Democracy [Boix et al.] * Post Cold War			0.42*	(0.24)
Constant	-1.63***	(0.24)	-1.70***	(0.25)
DV = WIN				
Number of Partners	0.075**	(0.030)	0.073**	(0.029)
Democracy [Polity IV]	0.010*	(0.0059)		
Democracy [Boix et al.]			0.19**	(0.088)
CINC score	1.55***	(0.43)	1.55***	(0.43)
Opponent(s)' CINC score	-0.76	(0.55)	-0.75	(0.55)
Troop Quality	0.012	(0.015)	0.0090	(0.015)
Dyad MID Propensity	5.83**	(2.28)	5.97***	(2.28)
System Level Democratic Share [Boix]	-1.34***	(0.51)	-1.34***	(0.52)
Post World War II	-0.71***	(0.13)	-0.73***	(0.13)
Post Cold-War	-0.11	(0.18)	-0.11	(0.18)
Partners * Post WWII	0.019	(0.037)	0.020	(0.037)
Partners * Post Cold War	0.030	(0.033)	0.030	(0.033)
Constant	-0.76***	(0.17)	-0.81***	(0.18)
Arc-Hyperbolic Tangent	0.086	(0.12)	0.092	(0.12)
Observations	4534		4509	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Time Sensitivity Analysis: Bivariate Probit and Bivariate Ordered Probit

	(1) Wars Only		(2) Wars Only		(3) MIDs 3-5		(4) MIDs 3-5	
DV = Ordered Partners								
Democracy [Polity IV]	0.012	(0.016)			0.014*	(0.0070)		
Democracy [Boix et al.]			0.44*	(0.27)			0.15	(0.096)
CINC score	-0.17	(1.34)	-0.94	(1.20)	1.91***	(0.39)	1.83***	(0.39)
Opponent(s)' CINC score	0.91	(1.03)	1.42	(1.04)	1.82***	(0.51)	1.76***	(0.51)
Dyad MID Propensity	3.24	(6.88)	9.50	(7.18)	2.03	(2.36)	2.09	(2.38)
Sys. Democratic Share	3.88**	(1.67)			0.77	(0.52)	0.79	(0.52)
Post World War II	-0.72*	(0.42)	-0.19	(0.36)	-0.31**	(0.13)	-0.32**	(0.13)
Post Cold-War	-0.64	(0.66)	0.12	(0.64)	-0.0091	(0.19)	-0.23	(0.19)
Polity * Post WWII	0.015	(0.033)			-0.011	(0.010)		
Polity * Post Cold War	0.064	(0.059)			0.059***	(0.018)		
Boix et al.* Post Cold War			0.088	(1.06)			0.64***	(0.23)
Boix et al. * Post WWII			0.013	(0.030)			-0.00081	(0.0084)
Cut 1	0.26	(0.42)	-0.19	(0.25)	0.92***	(0.16)	0.96***	(0.17)
Cut2	0.84**	(0.41)	0.34	(0.24)	1.49***	(0.16)	1.53***	(0.17)
DV = Victory								
Number of Partners	0.069**	(0.029)	0.063**	(0.027)	0.066**	(0.026)	0.064**	(0.026)
Democracy [Polity IV]	0.023	(0.015)			0.011***	(0.0036)		
Democracy [Boix et al.]			0.53**	(0.24)			0.20***	(0.057)
CINC score	5.12***	(1.80)	5.05***	(1.80)	2.22***	(0.45)	2.18***	(0.45)
Opponent(s)' CINC score	-3.33***	(0.87)	-3.28***	(0.86)	-2.46***	(0.55)	-2.46***	(0.55)
Troop Quality	0.066	(0.078)	0.044	(0.079)	0.023*	(0.014)	0.021	(0.014)
Dyad MID Propensity	-2.94	(8.00)	-3.23	(8.05)	-1.44	(1.70)	-1.29	(1.70)
Sys. Democratic Share	0.0041	(2.12)	0.10	(2.03)	-0.064	(0.54)	-0.058	(0.54)
Post World War II	0.29	(0.49)	0.16	(0.49)	0.0073	(0.096)	-0.021	(0.098)
Post Cold-War	-0.16	(0.55)	-0.095	(0.51)	-0.15	(0.11)	-0.15	(0.11)
Partners * Post WWII	-0.059	(0.037)	-0.053	(0.036)	-0.0045	(0.035)	-0.0034	(0.035)
Partners * Post C. War	0.018	(0.039)	0.017	(0.040)	0.014	(0.029)	0.015	(0.029)
Cut 1	-0.30	(0.61)	-0.20	(0.62)	-1.29***	(0.18)	-1.24***	(0.18)
Cut 2	0.42	(0.61)	0.54	(0.62)	1.16***	(0.18)	1.22***	(0.18)
Arc-Hyperbolic Tangent	0.018	(0.20)	0.016	(0.19)	0.0037	(0.061)	0.0084	(0.062)
Observations	336		334		4104		4080	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.7 Controlling for Initiation and its Interaction with Regime Type

One of the arguments made by Reiter and Stam (1998, 2002) is that democracies are particularly likely to win the wars they initiate. As we note in the body of the paper, at the margin, it is possible for both Reiter and Stam to be correct and for us to be correct. Our argument regarding democratic victory functions both when democracies initiate conflicts and when they are targeted. We use this section simply to show that our coalition size argument for democratic victory still obtains when controlling both for whether the state in question initiates the conflict, and controlling for the interaction effect of regime type and initiation. Our results are, in most models, robust to the inclusion of these additional variables, though in some specifications we lose statistical significance in models limited to wars only.

Overall, we continue to find strong evidence that democracies fight alongside larger coalitions, that states fighting alongside larger coalitions are more likely to win, and that democracies are more likely to win *because* they fight alongside larger coalitions. However, we also find evidence (Tables 20 and 21) consistent with Reiter and Stam's (1998, 2002) argument: to the extent that there is a direct effect of regime type on victory, this effect is stronger in conflicts that are initiated by the state in question. Democracies are most likely to win the conflicts they initiate. Thus, an important role of this section is to highlight the point that our theory (and results) can co-exist well with Reiter and Stam's work. We have identified a particularly strong channel through which regime type affects victory – i.e. through coalition size – but issues related to the what type of conflicts democracies initiate may also factor in at the margin.

Table 18 (parallel to Table 1 in the body of the paper) shows that democracies fight alongside more coalition partners than autocracies, and that they do so both when they initiate conflict and when they do not. Indeed, we find no significant relationship between initiation and coalition size.

Table 18: Number of Partners: Controlling for Initiation and Its Interaction with Democracy

	Sample = Wars Only				Sample = MIDs (3-5)			
Democracy [Polity IV]	0.038***	0.042***			0.060***	0.055**		
	(0.0088)	(0.011)			(0.015)	(0.023)		
Democracy [Boix et al.]			0.82***	0.86***			1.08***	1.03***
			(0.15)	(0.20)			(0.21)	(0.31)
Initiator	-0.17	-0.18	-0.16		-0.29	-0.29	-0.25	-0.28
	(0.34)	(0.34)	(0.32)		(0.35)	(0.35)	(0.34)	(0.29)
Polity * Initiator		-0.0076				0.010		
		(0.017)				(0.029)		
Democracy * Initiator				-0.068				0.095
				(0.32)				(0.44)
CINC score	0.89	0.90	0.32	0.33	-1.39	-1.44	-1.43	-1.45
	(0.98)	(0.99)	(0.95)	(0.97)	(1.07)	(1.05)	(1.02)	(1.00)
Opponent(s)' CINC score	0.80	0.77	1.04	0.92	2.17	2.22	2.43*	2.45*
	(1.12)	(1.11)	(1.05)	(1.18)	(1.38)	(1.38)	(1.39)	(1.39)
Constant	1.63***	1.64***	1.33***	1.27***	0.33	0.33	-0.11	-0.098
	(0.30)	(0.30)	(0.28)	(0.22)	(0.38)	(0.39)	(0.34)	(0.31)
Constant	0.50**	0.50**	0.49**	0.49**	2.10***	2.10***	2.09***	2.09***
	(0.22)	(0.22)	(0.22)	(0.22)	(0.093)	(0.094)	(0.092)	(0.092)
Observations	397	397	410	410	5019	5019	5229	5229

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$
Specification: Negative Binomial regression with errors clustered on dispute-side.

Table 19 (parallel to Table 2 in the paper) shows that the estimated effect of coalition size on the probability of victory remains positive when controlling for the effects of regime type, initiation, and the interaction of regime type and initiation. However, in models where the sample is restricted to wars only, these results fall from statistical significance. It is notable that states that initiate conflict are, *ceteris paribus* less likely to prevail in that conflict.

Table 19: Probability of Victory: Controlling for Initiation and its Interaction with Democracy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample = Wars Only				Sample = MIDs (3-5)			
Number of Partners	0.035*	0.036*	0.031	0.032	0.068***	0.068***	0.069***	0.069***
	(0.019)	(0.018)	(0.020)	(0.019)	(0.013)	(0.013)	(0.013)	(0.013)
Democracy [Polity IV]	0.011	0.0030			0.0091**	0.0076		
	(0.015)	(0.021)			(0.0037)	(0.0051)		
Democracy [Boix et al.]			0.37	0.16			0.16***	0.071
			(0.25)	(0.30)			(0.057)	(0.081)
Initiator	-0.76***	-0.73**	-0.81***	-0.97***	-0.26***	-0.26***	-0.26***	-0.32***
	(0.29)	(0.30)	(0.29)	(0.30)	(0.067)	(0.067)	(0.067)	(0.080)
Polity * Initiator		0.020				0.0030		
		(0.029)				(0.0076)		
Democracy * Initiator				0.62				0.18
				(0.45)				(0.12)
CINC score	5.50***	5.62***	5.30***	5.59***	2.57***	2.57***	2.62***	2.63***
	(1.73)	(1.68)	(1.69)	(1.64)	(0.46)	(0.46)	(0.44)	(0.44)
Opponent(s)' CINC score	-2.70***	-2.71***	-2.59***	-2.60***	-2.58***	-2.58***	-2.59***	-2.59***
	(0.86)	(0.88)	(0.86)	(0.88)	(0.53)	(0.53)	(0.53)	(0.52)
Dyad MID Propensity	-4.73	-4.07	-4.66	-3.48	-1.64	-1.61	-1.33	-1.23
	(7.47)	(7.46)	(7.36)	(7.31)	(1.75)	(1.75)	(1.73)	(1.73)
Troop Quality	0.097	0.085	0.074	0.050	0.024*	0.024*	0.017	0.016
	(0.084)	(0.080)	(0.084)	(0.081)	(0.014)	(0.014)	(0.011)	(0.011)
Cut 1: Constant	-0.72**	-0.71**	-0.68**	-0.73**	-1.38***	-1.38***	-1.32***	-1.36***
	(0.31)	(0.31)	(0.33)	(0.33)	(0.072)	(0.072)	(0.078)	(0.082)
Cut 2: Constant	0.048	0.061	0.10	0.060	1.09***	1.09***	1.15***	1.12***
	(0.34)	(0.34)	(0.33)	(0.33)	(0.073)	(0.073)	(0.076)	(0.079)
Observations	336	336	345	345	4104	4104	4234	4234

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Ordered Probit with errors clustered on dispute-side.

Table 20 shows the bivariate probit results (parallel to Table 5 in the paper), controlling for initiation and the interaction between initiation and regime type in both stages. Again, our results are generally robust, though the effect of coalition size on victory is not statistically significant in the wars-only sample. We see that democracies are more likely to fight in larger coalitions and that states fighting in larger coalitions are more likely to win.

Table 20: Joint Probability of Partners and Victory: Bivariate Probit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample = Wars Only				Sample = MIDs (3-5)			
DV = Coalition								
Democracy [Polity IV]	0.047***	0.053***			0.030***	0.024**		
	(0.015)	(0.020)			(0.0076)	(0.011)		
Democracy [Boix et al.]			0.67***	0.76**			0.46***	0.37**
			(0.24)	(0.31)			(0.11)	(0.15)
Initiator	-0.0027	-0.019	-0.083	-0.021	0.12	0.10	0.11	0.038
	(0.32)	(0.32)	(0.33)	(0.38)	(0.16)	(0.15)	(0.16)	(0.15)
Polity * Initiator		-0.014				0.013		
		(0.032)				(0.014)		
Democracy * Initiator				-0.23				0.17
				(0.50)				(0.20)
CINC score	-0.53	-0.51	-0.55	-0.55	-0.41	-0.43	-0.32	-0.33
	(1.47)	(1.47)	(1.42)	(1.42)	(0.60)	(0.60)	(0.59)	(0.59)
Opponent(s)' CINC score	1.49	1.48	1.45	1.44	2.06***	2.06***	2.08***	2.07***
	(1.02)	(1.02)	(1.04)	(1.03)	(0.68)	(0.68)	(0.68)	(0.68)
Dyad MID Propensity	15.7*	15.1	15.9*	15.4	3.15	3.28	3.75	3.85
	(9.31)	(9.26)	(9.36)	(9.36)	(3.26)	(3.27)	(3.25)	(3.25)
Constant	-0.47	-0.45	-0.66*	-0.67*	-1.45***	-1.44***	-1.63***	-1.60***
	(0.37)	(0.37)	(0.38)	(0.37)	(0.17)	(0.17)	(0.16)	(0.16)
DV = Win								
Number of Partners	0.050*	0.050	0.047	0.044	0.089***	0.089***	0.089***	0.089***
	(0.030)	(0.030)	(0.031)	(0.032)	(0.018)	(0.017)	(0.018)	(0.017)
Democracy [Polity IV]	0.00064	-0.023			-0.0013	-0.014*		
	(0.021)	(0.028)			(0.0057)	(0.0079)		
Democracy [Boix et al.]			0.13	-0.18			-0.081	-0.30**
			(0.34)	(0.42)			(0.088)	(0.13)
Initiator	-0.82**	-0.78**	-0.84**	-1.16***	-0.16*	-0.18**	-0.17*	-0.36***
	(0.35)	(0.35)	(0.36)	(0.32)	(0.091)	(0.090)	(0.090)	(0.10)
Polity * Initiator		0.064**				0.028***		
		(0.028)				(0.0094)		
Democracy * Initiator				1.02**				0.48***
				(0.45)				(0.15)
CINC score	5.46***	5.86***	5.26***	5.78***	3.45***	3.47***	3.60***	3.65***
	(1.49)	(1.47)	(1.45)	(1.44)	(0.43)	(0.43)	(0.40)	(0.40)
Opponent(s)' CINC score	-2.32**	-2.35**	-2.20**	-2.19**	-0.16	-0.15	-0.22	-0.22
	(0.92)	(0.94)	(0.92)	(0.96)	(0.52)	(0.52)	(0.52)	(0.51)
Troop Quality	0.14	0.11	0.13	0.098	0.033*	0.031*	0.026*	0.024*
	(0.085)	(0.084)	(0.083)	(0.083)	(0.017)	(0.017)	(0.013)	(0.013)
Dyad MID Propensity	-7.14	-5.04	-6.39	-4.06	2.59	2.84	2.78	3.04
	(8.52)	(8.36)	(8.63)	(8.66)	(2.39)	(2.37)	(2.36)	(2.34)
Constant	-0.17	-0.20	-0.19	-0.12	-1.58***	-1.57***	-1.53***	-1.46***
	(0.34)	(0.34)	(0.34)	(0.33)	(0.095)	(0.094)	(0.098)	(0.10)
Arc-Hyperbolic Tangent	-0.22	-0.20	-0.21	-0.19	0.10	0.099	0.12	0.11
	(0.26)	(0.26)	(0.27)	(0.28)	(0.12)	(0.12)	(0.12)	(0.12)
Observations	344	344	353	353	4534	4534	4675	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To explore this a little further, in Table 21 we rerun the bivariate probit model, but using partners' summed CINC scores rather then the raw count of coalition partners. Here we see results that are very similar to those in Table 20, but the effect of coalition size on victory is statistically significant in wars as well as in MIDs. This helps extend our confidence in the strength of our core result.

As noted above, the results in both Table 20 and Table 21 are also consistent with a core argument of Reiter and Stam (1998, 2002). We see that the direct effect of regime type on victory is stronger in conflicts that are initiated by the state in question: democratic initiators are particularly likely to win.

Table 21: Joint Probability of Partners and Victory: Bivariate Probit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample = Wars Only				Sample = MIDs (3-5)			
DV = Coalition (CINC)								
Democracy [Polity IV]	0.10***	0.12***			0.045***	0.066***		
	(0.024)	(0.028)			(0.010)	(0.015)		
Democracy [Boix et al.]			1.66***	1.78***			0.59***	1.00***
			(0.37)	(0.45)			(0.15)	(0.26)
Initiator	-0.65	-0.57	-0.67*	-0.62	0.23	0.38**	0.24	0.65**
	(0.40)	(0.39)	(0.38)	(0.46)	(0.19)	(0.20)	(0.19)	(0.26)
Polity * Initiator		-0.039				-0.033**		
		(0.034)				(0.016)		
Democracy * Initiator				-0.26				-0.64**
				(0.47)				(0.28)
CINC score	-1.03	-1.19	-1.40	-1.66	0.57	0.58	0.75	0.75
	(2.41)	(2.50)	(2.33)	(2.46)	(0.82)	(0.82)	(0.79)	(0.79)
Opponent(s)' CINC score	0.93	0.96	1.11	1.19	1.15*	1.17*	1.18*	1.22*
	(0.95)	(0.93)	(0.97)	(0.95)	(0.63)	(0.64)	(0.63)	(0.66)
Dyad MID Propensity	-5.16	-7.54	-6.31	-8.41	0.30	-0.059	0.23	-0.29
	(9.79)	(9.18)	(10.2)	(9.59)	(4.66)	(4.61)	(4.64)	(4.59)
Constant	-1.14***	-1.10***	-1.70***	-1.68***	-2.31***	-2.43***	-2.52***	-2.80***
	(0.36)	(0.36)	(0.42)	(0.44)	(0.15)	(0.17)	(0.16)	(0.26)
DV = Win								
Partner(s)' CINC score	2.96**	3.89***	2.78**	3.81***	4.80***	4.70***	4.85***	4.79***
	(1.34)	(1.18)	(1.34)	(1.20)	(0.78)	(0.80)	(0.76)	(0.79)
Democracy [Polity IV]	-0.014	-0.054*			0.00022	-0.0092		
	(0.021)	(0.028)			(0.0057)	(0.0081)		
Democracy [Boix et al.]			-0.11	-0.72*			-0.033	-0.22*
			(0.37)	(0.43)			(0.085)	(0.13)
Initiator	-0.69**	-0.57*	-0.72**	-1.11***	-0.25**	-0.27**	-0.26**	-0.43***
	(0.33)	(0.30)	(0.33)	(0.31)	(0.11)	(0.11)	(0.11)	(0.10)
Polity * Initiator		0.088***				0.022**		
		(0.029)				(0.011)		
Democracy * Initiator				1.46***				0.43**
				(0.46)				(0.17)
CINC score	5.08***	5.44***	4.88***	5.47***	2.53***	2.55***	2.72***	2.77***
	(1.27)	(1.29)	(1.23)	(1.30)	(0.51)	(0.52)	(0.49)	(0.50)
Opponent(s)' CINC score	-2.51***	-2.63***	-2.33***	-2.45***	-0.29	-0.26	-0.34	-0.33
	(0.77)	(0.75)	(0.74)	(0.74)	(0.54)	(0.54)	(0.53)	(0.53)
Troop Quality	0.16**	0.10	0.15**	0.084	0.048***	0.046***	0.036**	0.034**
	(0.067)	(0.066)	(0.065)	(0.067)	(0.017)	(0.017)	(0.014)	(0.014)
Dyad MID Propensity	-2.58	-1.06	-2.22	-0.26	4.53*	4.74**	4.46*	4.73**
	(8.15)	(7.85)	(8.14)	(7.98)	(2.39)	(2.41)	(2.33)	(2.36)
Constant	-0.31	-0.38	-0.23	-0.14	-1.54***	-1.54***	-1.51***	-1.44***
	(0.31)	(0.31)	(0.32)	(0.32)	(0.11)	(0.12)	(0.11)	(0.11)
Arc-Hyperbolic Tangent	-0.44*	-0.67**	-0.48*	-0.77**	-0.15	-0.13	-0.16	-0.14
	(0.24)	(0.28)	(0.26)	(0.35)	(0.16)	(0.16)	(0.16)	(0.16)
Observations	344	344	353	353	4534	4534	4675	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.8 Testing Robustness to Reiter, Stam, and Horowitz's Revised War Data

Here we demonstrate that our results are robust to a brand new dataset for Wars from 1816-2007 developed by Reiter, Stam and Horowitz (2014). We refer to this dataset as the RSH data below. The RSH data cover wars only and not MIDs, which is somewhat limiting, but the data still offer a valuable opportunity to demonstrate the robustness of our results to alternative codings of coalition size and victory.

The RSH data represent an exciting new step forward in the quantitative study of interstate war. These data are built on the backbone of the original Correlates of War dataset (which, in turn, forms the basis of the Maoz Dyadic MID dataset we use in our paper). What Reiter, Stam and Horowitz do is to recode a number of conflicts, breaking larger conflicts apart into their constituent pieces and, essentially, making different coding choices on a number of cases where reasonable scholars might disagree. Thus, the RSH data offers scholars an alternative conflict dataset on which to test results. If results are robust to both to estimation using the original COW (Maoz) data and to estimation using this new dataset, then scholars can be confident the results are not driven by errors or controversial coding choices in either dataset. We exploit that opportunity here.

Tables 22 (parallel to Tables 1 and 3 in the paper), 23 (parallel to Tables 2 and 4 in the paper), and 24 (parallel to tables 5 and 6 in the paper) replicate our core results with this alternative dataset. Note that these results are all "wars only" as the RSH data does not cover non-war MIDs. Figures 2-6 show these same results graphically.

Table 22 and Figures 2 and 3 show a strong positive effect of democracy on coalition size. Table 23 and Figures 4, 5, and 6 show a strong positive effect of coalition size on the probability of victory. When we combine these analyses in a bivariate probit setup (Table 24), we see that democracies win the wars they fight because they fight in larger coalitions. Note that the one departure from robustness in this data involves models 1 and 2 in Table 21. In these models, while the estimated effect of regime type on coalition size is positive and strong, the effect of coalition size on victory is statistically zero. However, in models 3 and 4 where we operationalize coalition size as partners' summed CINC score instead of as the number of

partners, a strong positive effect of coalition size on victory is once again observed. Thus, the overall effect of this supplementary analysis with the RSH data is to increase our confidence that the results we report in the body of our paper reflect genuine real-world relationships and are not driven by errors or peculiarities of coding within the MIDs data.

We further explore these new data in the permutation and cross validation tests below.

Table 22: Number of Coalition Partners: RSH Wars

	DV=Count	DV=Count	DV=Power	DV=Power
Democracy [Polity IV]	0.054***		0.0066***	
	(0.015)		(0.0018)	
Democracy [Boix et al.]		1.34***		0.13***
		(0.20)		(0.031)
CINC score	-1.66	-2.31*	-0.19	-0.20
	(1.39)	(1.38)	(0.17)	(0.16)
Opponent(s)' CINC score	-0.41	-0.54	0.18	0.15
	(1.88)	(1.75)	(0.13)	(0.12)
Constant	0.94**	0.46	0.083***	0.042**
	(0.46)	(0.43)	(0.024)	(0.018)
Constant	1.20***	1.07***		
	(0.22)	(0.23)		
Observations	402	435	402	435
R^2			0.117	0.156

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Models 1 & 2: Negative Binomial regression with errors clustered on dispute-side.

Models 3 & 4: OLS with errors clustered on dispute-side.

Table 23: Probability of Victory: RSH Wars

	(1) Count	(2) Count	(3) Power	(4) Power
Number of Partners	0.072***	0.067**		
	(0.027)	(0.027)		
Partner(s)' CINC score			3.17***	3.06***
			(0.80)	(0.74)
Democracy [Polity IV]	0.041***		0.032**	
	(0.013)		(0.014)	
Democracy [Boix et al.]		0.82***		0.69***
		(0.21)		(0.23)
CINC score	2.79	3.15*	3.18*	3.61**
	(1.71)	(1.64)	(1.79)	(1.71)
Opponent(s)' CINC score	-4.38***	-4.50***	-4.86***	-4.97***
	(0.93)	(0.95)	(1.11)	(1.11)
Dyad MID Propensity	-0.98	-2.70	-1.38	-3.15
	(5.30)	(4.85)	(5.50)	(4.96)
Troop Quality	0.070	0.039	0.084	0.050
	(0.077)	(0.061)	(0.074)	(0.063)
Cut 1: Constant	-0.38**	-0.15	-0.31	-0.12
	(0.19)	(0.21)	(0.20)	(0.21)
Cut 2: Constant	0.014	0.26	0.10	0.31*
	(0.17)	(0.19)	(0.17)	(0.18)
Observations	344	360	344	360

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Specification: Ordered Probit with errors clustered on dispute-side.

Table 24: Joint Probability of Partners and Victory: Bivariate Probit

	(1) Partners	(2) Partners	(3) Power	(4) Power
ALLY				
Democracy [Polity IV]	0.055***		0.071***	
	(0.015)		(0.018)	
Democracy [Boix et al.]		1.10***		1.06***
		(0.22)		(0.21)
CINC score	0.34	-0.056	-3.04	-2.86
	(1.60)	(1.57)	(2.75)	(2.77)
Opponent(s)' CINC score	-0.89	-0.94	1.69	1.53
	(1.10)	(1.06)	(1.20)	(1.21)
Dyad MID Propensity	10.2	9.62	14.5*	11.5
	(6.74)	(6.58)	(7.60)	(7.81)
Constant	-0.67**	-1.01***	-1.65***	-1.94***
	(0.28)	(0.28)	(0.27)	(0.26)
WIN				
Number of Partners	-0.015	-0.020		
	(0.056)	(0.055)		
Partner(s)' CINC score			3.47***	3.24**
			(1.18)	(1.27)
Democracy [Polity IV]	0.049***		0.024	
	(0.013)		(0.016)	
Democracy [Boix et al.]		0.97***		0.53*
		(0.22)		(0.29)
CINC score	2.94	3.42*	3.45*	3.96**
	(1.81)	(1.76)	(1.95)	(1.88)
Opponent(s)' CINC score	-4.82***	-4.93***	-5.36***	-5.48***
	(1.38)	(1.40)	(1.73)	(1.74)
Troop Quality	0.12	0.092	0.11	0.082
	(0.076)	(0.069)	(0.076)	(0.068)
Dyad MID Propensity	-3.72	-6.70	-6.23	-8.51
	(6.10)	(5.70)	(6.15)	(5.79)
Constant	0.21	-0.078	-0.057	-0.21
	(0.20)	(0.21)	(0.19)	(0.21)
Constant	0.38	0.36	-0.40	-0.31
	(0.26)	(0.24)	(0.30)	(0.36)
Observations	344	360	344	360

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

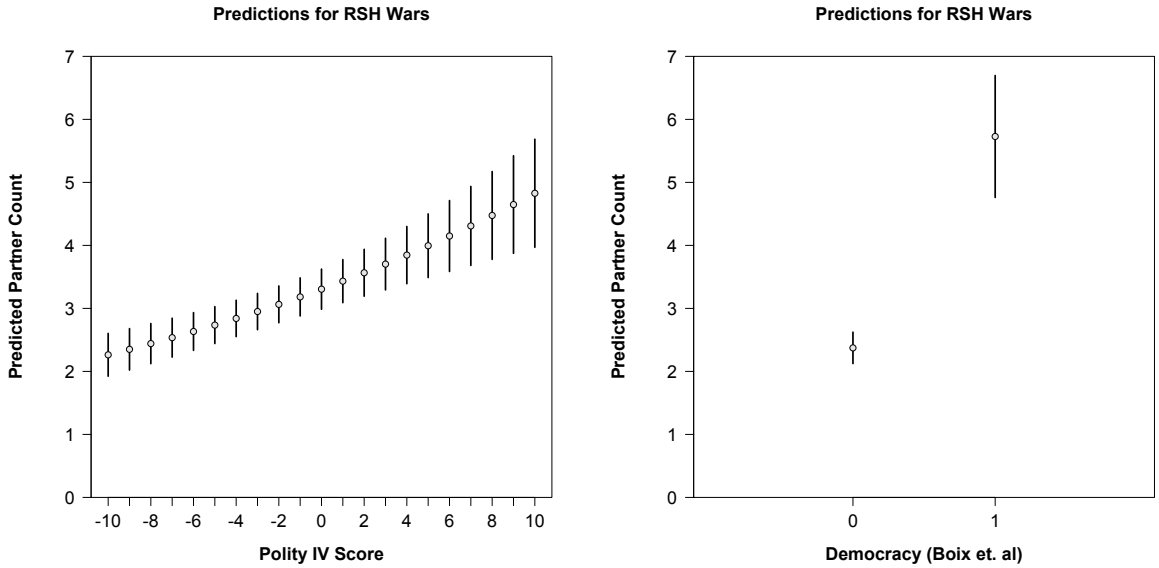

Figure 2: Predicted number of partners as a function of two different measures of democracy for the sample of Wars from the RSH dataset (Reiter, Stam and Horowitz, 2014). All other variables from the models are held at their mean values.

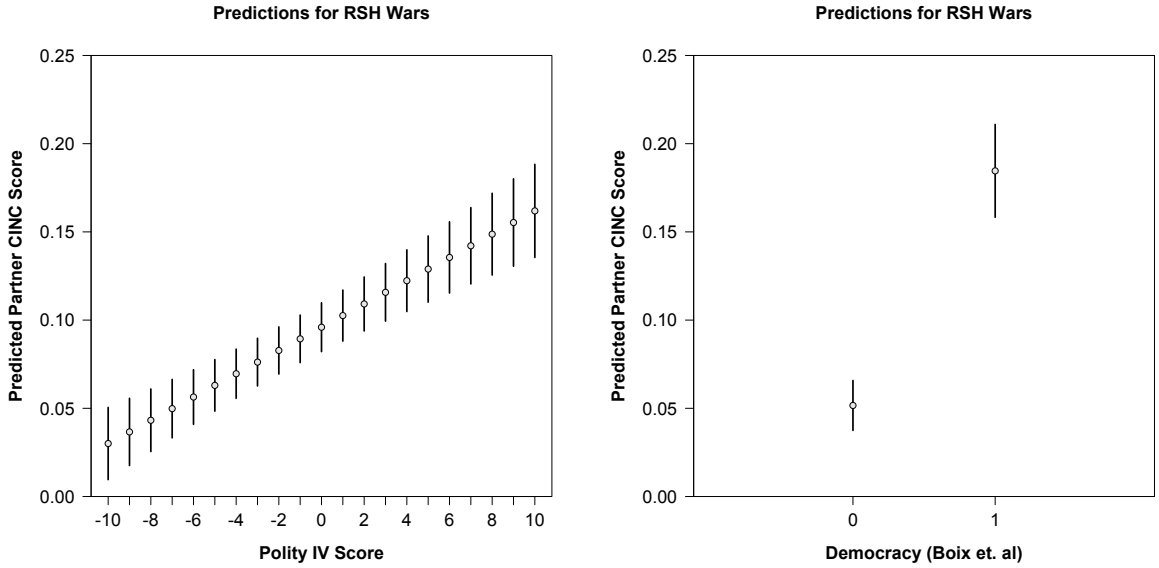

Figure 3: Predicted partners' summed CINC scores a function of two different measures of democracy for the sample of Wars from the RSH dataset (Reiter, Stam and Horowitz, 2014). All other variables from the models are held at their mean values.

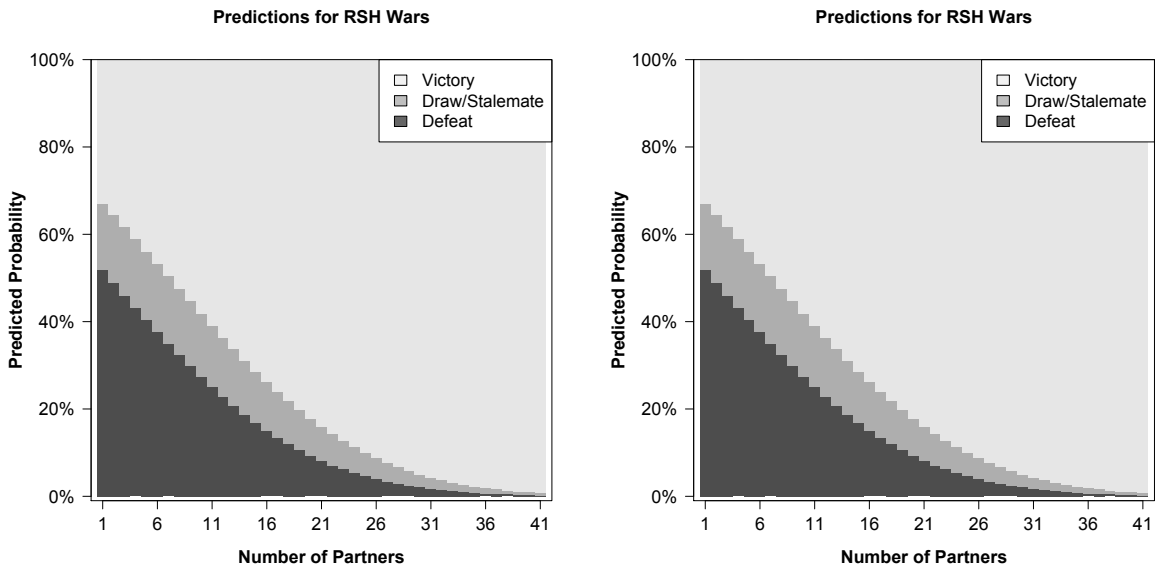

Figure 4: Predicted probability of ordered outcomes during RSH Wars as a function of the number of coalition partners using two different measures of democracy (Polity IV in the two left panels and Boix, Miller and Rosato (2013) in the two right panels). All other variables from the models are held at their mean or median values. The war data is from Reiter, Stam and Horowitz (2014).

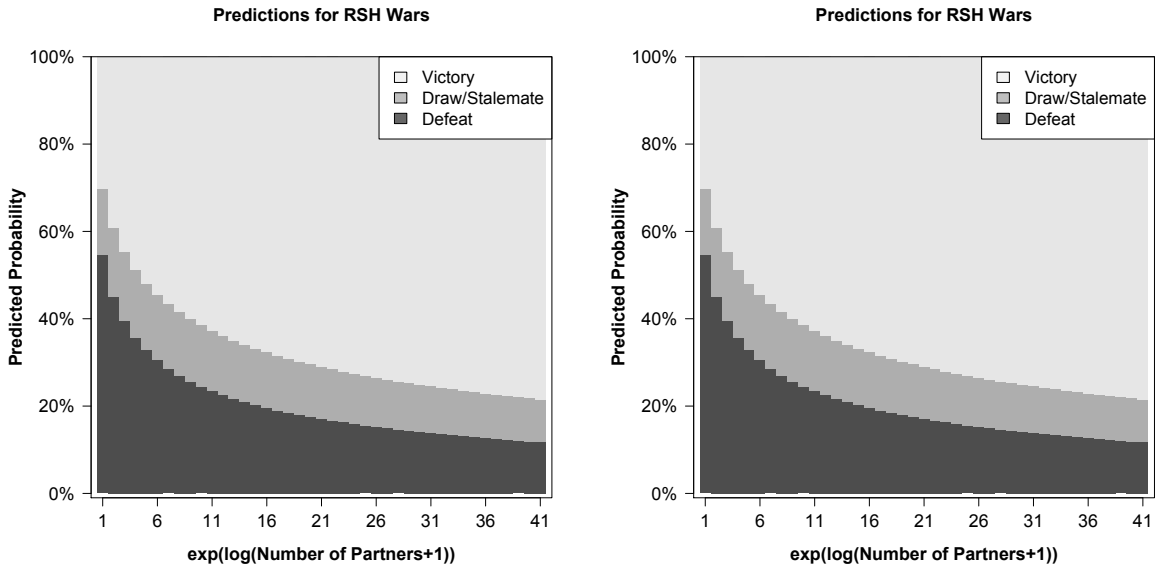

Figure 5: Predicted probability of ordered outcomes during RSH Wars as a function of the logged number of coalition partners using two different measures of democracy (Polity IV in the two left panels and Boix, Miller and Rosato (2013) in the two right panels). All other variables from the models are held at their mean or median values. The war data is from Reiter, Stam and Horowitz (2014).

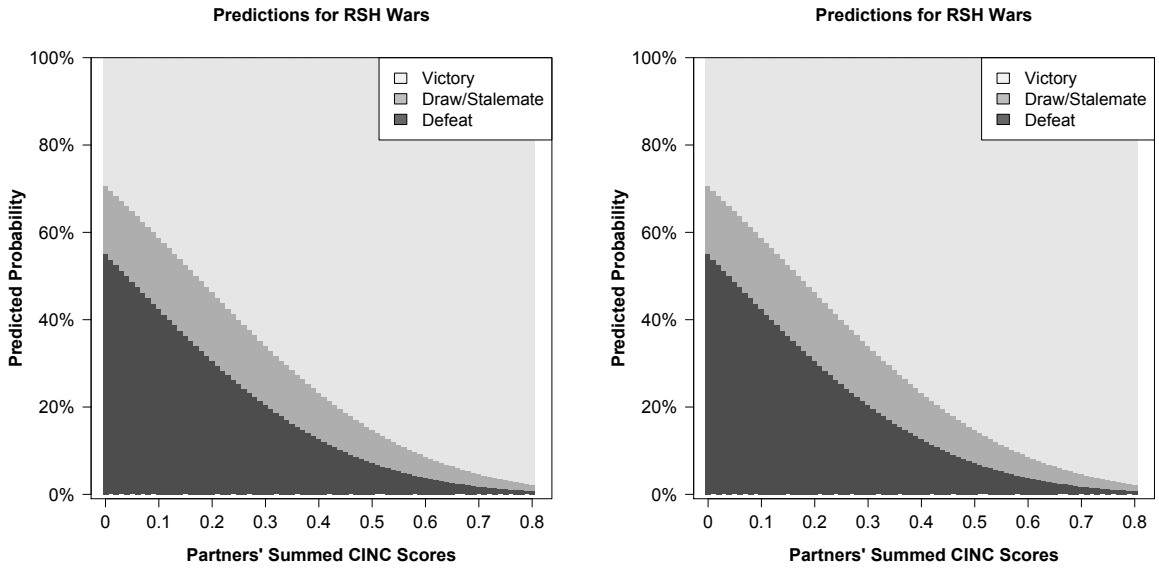

Figure 6: Predicted probability of ordered outcomes during RSH Wars as a function of partners' summed CINC scores using two different measures of democracy (Polity IV in the two left panels and Boix, Miller and Rosato (2013) in the two right panels). All other variables from the models are held at their mean or median values. The war data is from Reiter, Stam and Horowitz (2014).

10.9 Permutation Analysis for Ordered Probit Models

In the main paper, the second link of our argument (Hypothesis 2) is that states that are accompanied by more partners are more likely to prevail in both wars and MIDs more generally. To test this hypothesis, we estimate the probability of victory using ordered probit regression. In this section we conduct permutation tests (e.g., Gordon, 2005) in order to determine if these results are an artifact of chance instead of a product of our theoretical mechanism.

For the permutation test reported here, we seek to determine if the coefficient from the ordered probit models is different from a random re-ordering of the number of coalition partners in the data. To check for this possibility we ran 1,000 regressions for each of the models, randomly re-ordering the number of coalition partners variable for each regression, keeping the other variables fixed. The figures in this section present the null distribution and the observed coefficient from the original models reported in the manuscript. For each combination of observations (MIDS 3-5, Wars only (both from (Maoz, 2005)), and Wars from the RSH dataset (Reiter, Stam and Horowitz, 2014)) and different operationalizations of the number of partners (logged partner count, unlogged partner count, and the partners' summed CINC scores) and two different measures of democracy (Polity IV, and Boix, Miller and Rosato (2013)). Each of the observed coefficients for the ordered probit models exist outside the null distributions as seen in the figures below. The permutation tests provide strong support for the results reported in the paper. The results from these tests indicate that the fit of our models, the coefficients and standard errors, are not the result of random chance but instead represent a general pattern of state behavior and conflict outcomes in the international system.

10.9.1 Permutation of Number of Partners Variable

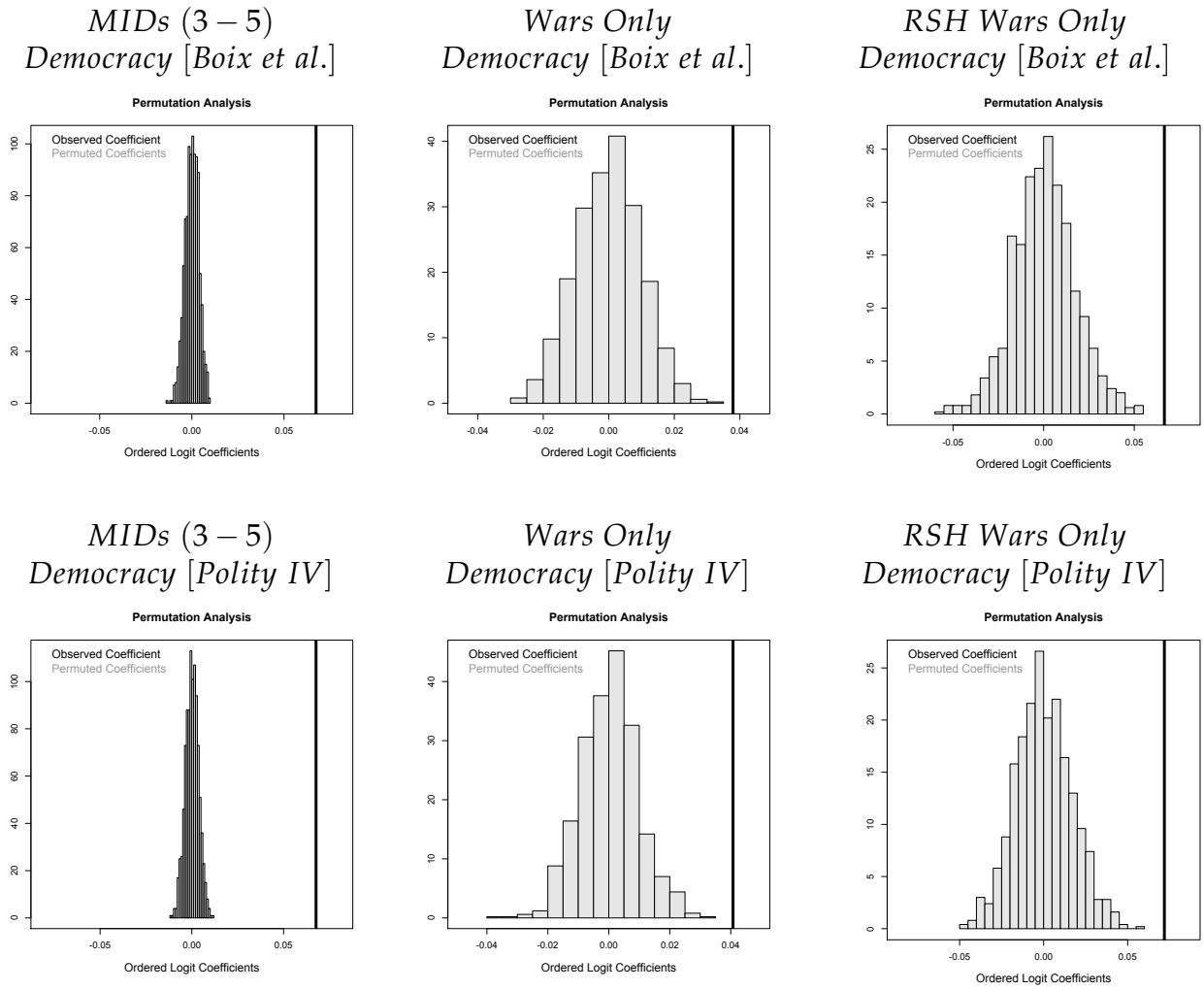

Figure 7: Distribution of 1,000 coefficients generated from ordered probit regression models using permuted collation counts (grey). This null distributions is compared to the coefficients generated from the un un-permuted data (black). The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.9.2 Permutation of Logged Number of Partners Variable

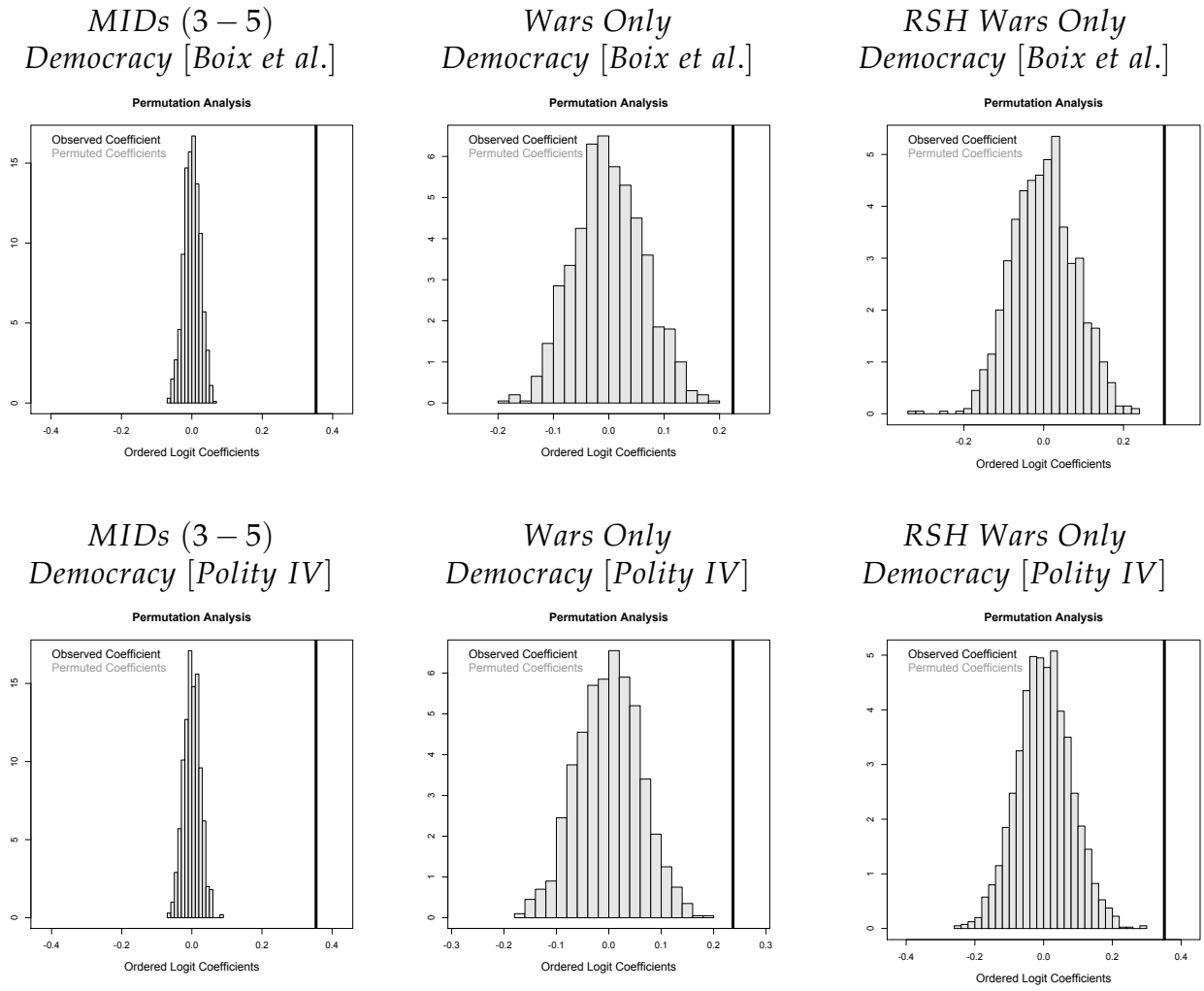

Figure 8: Distribution of 1,000 coefficients generated from ordered probit regression models using permuted collation counts (grey). This null distributions is compared to the coefficients generated from the un un-permuted data (black). The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.9.3 Permutation of Partners' Summed CINC Scores Variable

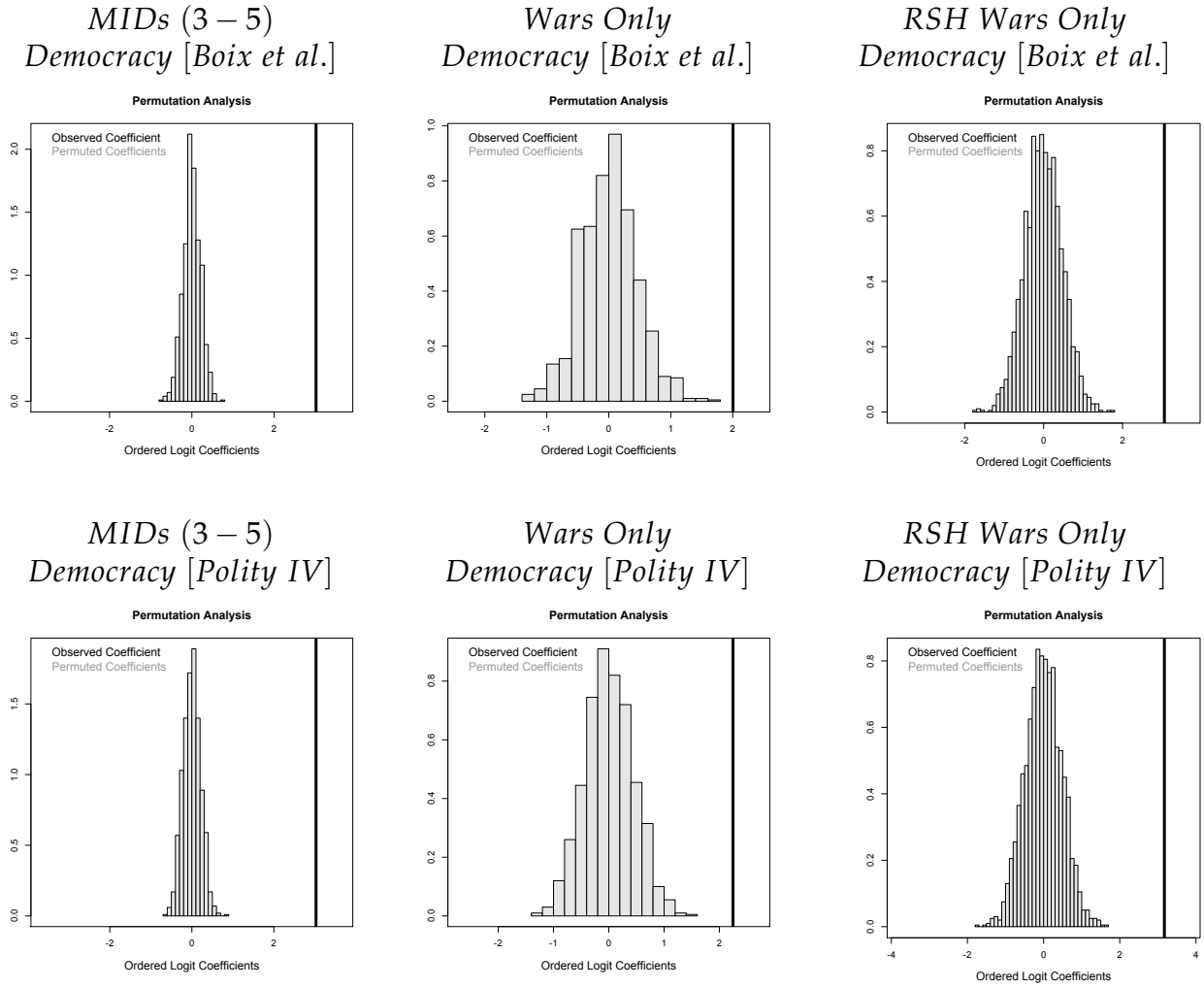

Figure 9: Distribution of 1,000 coefficients generated from ordered probit regression models using permuted collation counts (grey). This null distributions is compared to the coefficients generated from the un un-permuted data (black). The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.10 Cross Validation for Ordered Probit Models

Again, the second link of our argument (Hypothesis 2) is that states that are accompanied by more partners are more likely to prevail in both wars and MIDs more generally. Here we use cross validation methods, similar to those recommended by Ward, Greenhill and Bakke (2010), to evaluate the out of sample performance of the models that contain the coalition partners variable with a base-line model that only contains controls.[3] We follow Hill Jr. and Jones (2014) and estimate 1000 K-fold cross validated models for each combination of observations (MIDS 3-5, Wars only (both from Maoz (2005)), Wars from the RSH dataset (Reiter, Stam and Horowitz, 2014)) and different operationalizations of the number of partners (logged partner count, unlogged partner count, and the partners' summed CINC scores) and two different measures of democracy (Polity IV, and Boix, Miller and Rosato (2013)).

For each of the 1000 K-fold cross validated models we randomly divide the dataset into K=10 subsets. We then fit the model using the observations from 9 of the subsets of data and then predict the value of the ordered dependent variable for the 1 remaining out of sample data subset. We repeat this process for each of the 10 data subsets so that we predict a value of the dependent variable for every one of the original observations when they are in one of the out of sample data subsets.

Thus, for each of the 1000 simulations we are able to make an out of sample prediction $\hat{y}$ for all of the observations of the ordered dependent variables y. We then estimate the Spearman's ρ correlations coefficients (which can take values from -1 to 1) using the ordered dependent variable y and the predicted ordered variable $\hat{y}$. We use the Spearman's ρ because it is a non-parametric measure of association useful for modeling the relationship between ordered or ranked data. The closer the value of ρ is to 1 then more closely related the out of sample predictions $\hat{y}$ are the observed value y.

We use the same process for four distinct models, a baseline model with just controls, and then three additional models that include the same controls in addition to either the coalition

[3]For more information on the technical details of these techniques see work by Efron (1983), Geisser (1975), Hastie, Tibshirani and Friedman (2008), and Picard and Cook (1984). There are also a few examples of research in the political science literature that use these techniques (Hoff and Ward, 2004; Ward and Hoff, 2007; Ward, Siverson and Cao, 2007).

partner variable, the democracy variable, or both of these variables. We then compare the distribution of differences between ρ coefficients of these three models relative to the baseline model. The larger the difference the greater the improvement of the predictive power of the alternative model. In almost every model, the coalition variable substantially increases the predictive power of the model relative to the baseline. The democracy variable increase the predictive power in some models.

Overall, both variables in combination increase the predictive power of the model the most in all of the different combinations of observations (sample of MIDS 3-5, Wars only, and RHS Wars) and different operationalizations of the number of partners (logged partner count, unlogged partner count, and the partners' summed CINC scores) and two different measures of democracy (Polity IV, and Boix, Miller and Rosato (2013)). Just like the permutation tests, these cross validation models provide strong support for the results reported in the paper because models that contain the coalition variable always do a better job a predicting war outcomes than models that do not contain this variable. The models with both measures of democracy and coalition size do the better for some dependent variable because, as we argue in the paper, these variables are capturing overlapping concepts. Democracies go to war with more coalition partners and more coalition partners are associated with victory. Interestingly, for the models that use the RHS War data, the model with only the democracy variable does a worse job of predicting these outcomes than the baseline model.

10.10.1 Cross Validation of Number of Partners Variable

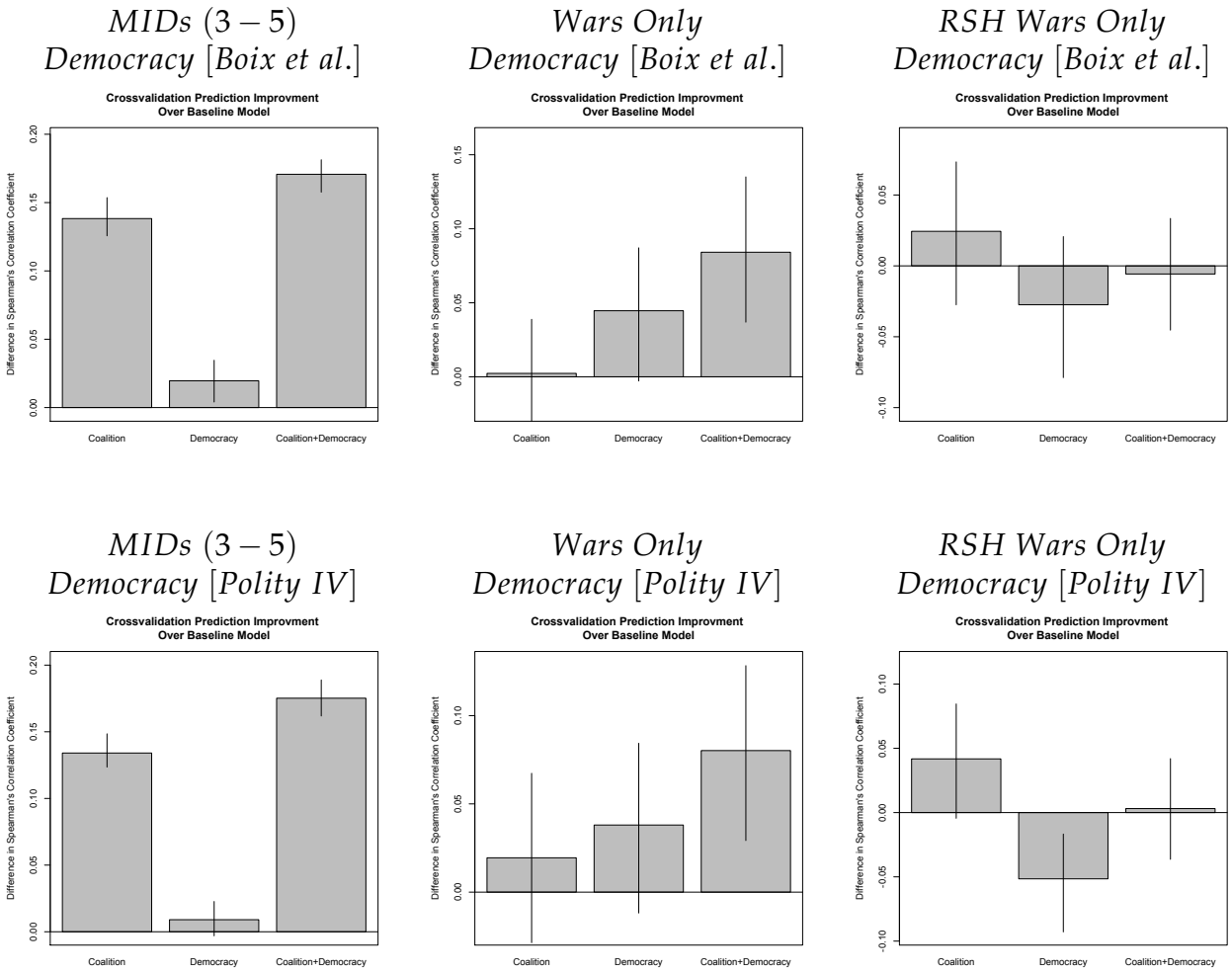

Figure 10: Difference of Spearman's ρ coefficients for three models compared to a baseline model (controls only). The ρ coefficients are correlations between observed and predicted ordered dependent variables. The larger the difference, the greater the improvement in the predictive power of the alternative model relative to the baseline. The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.10.2 Cross Validation of Logged Number of Partners Variable

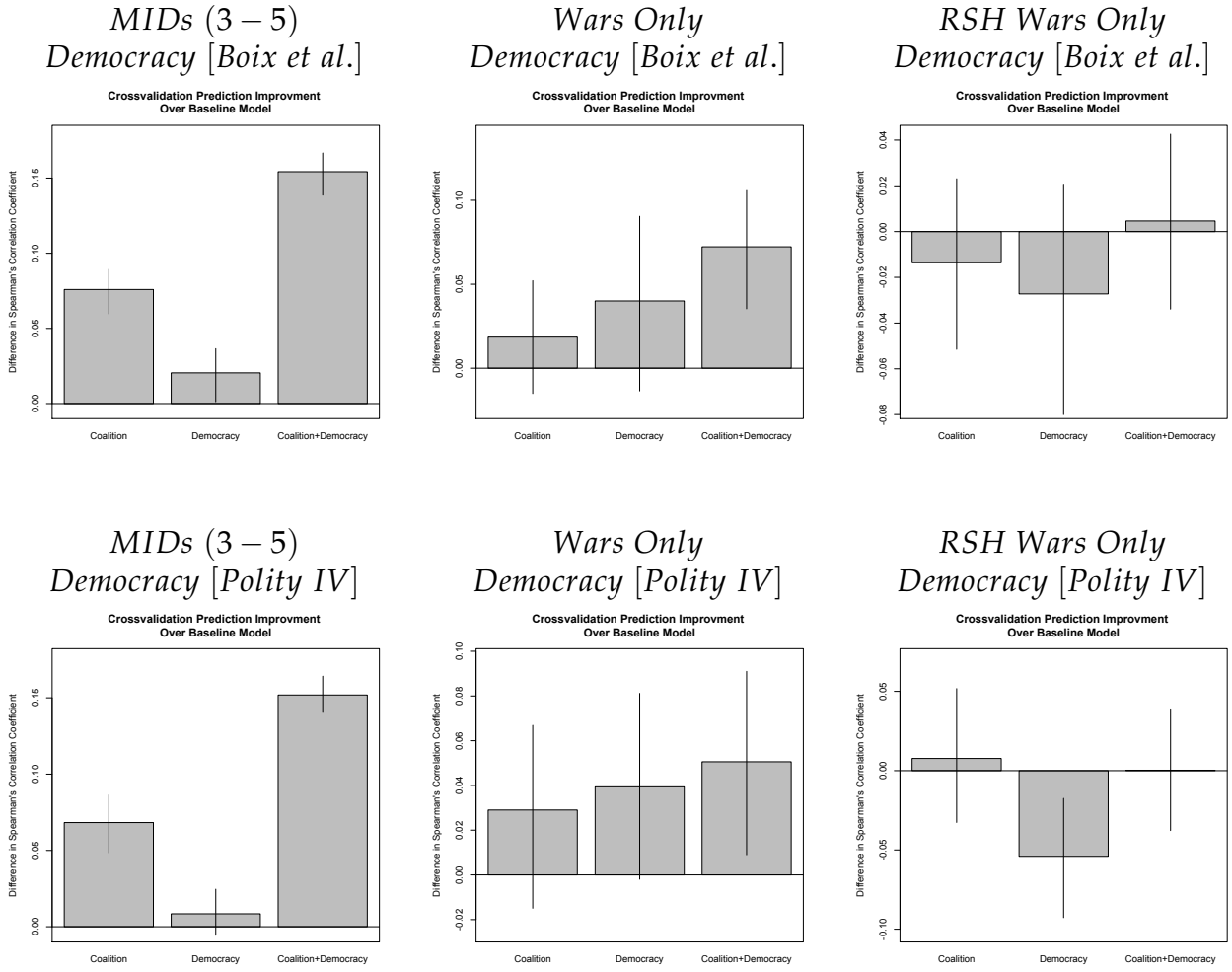

Figure 11: Difference of Spearman's ρ coefficients for three models compared to a baseline model (controls only). The ρ coefficients are correlations between observed and predicted ordered dependent variables. The larger the difference, the greater the improvement in the predictive power of the alternative model relative to the baseline. The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.10.3 Cross Validation of Partners' Summed CINC Scores Variable

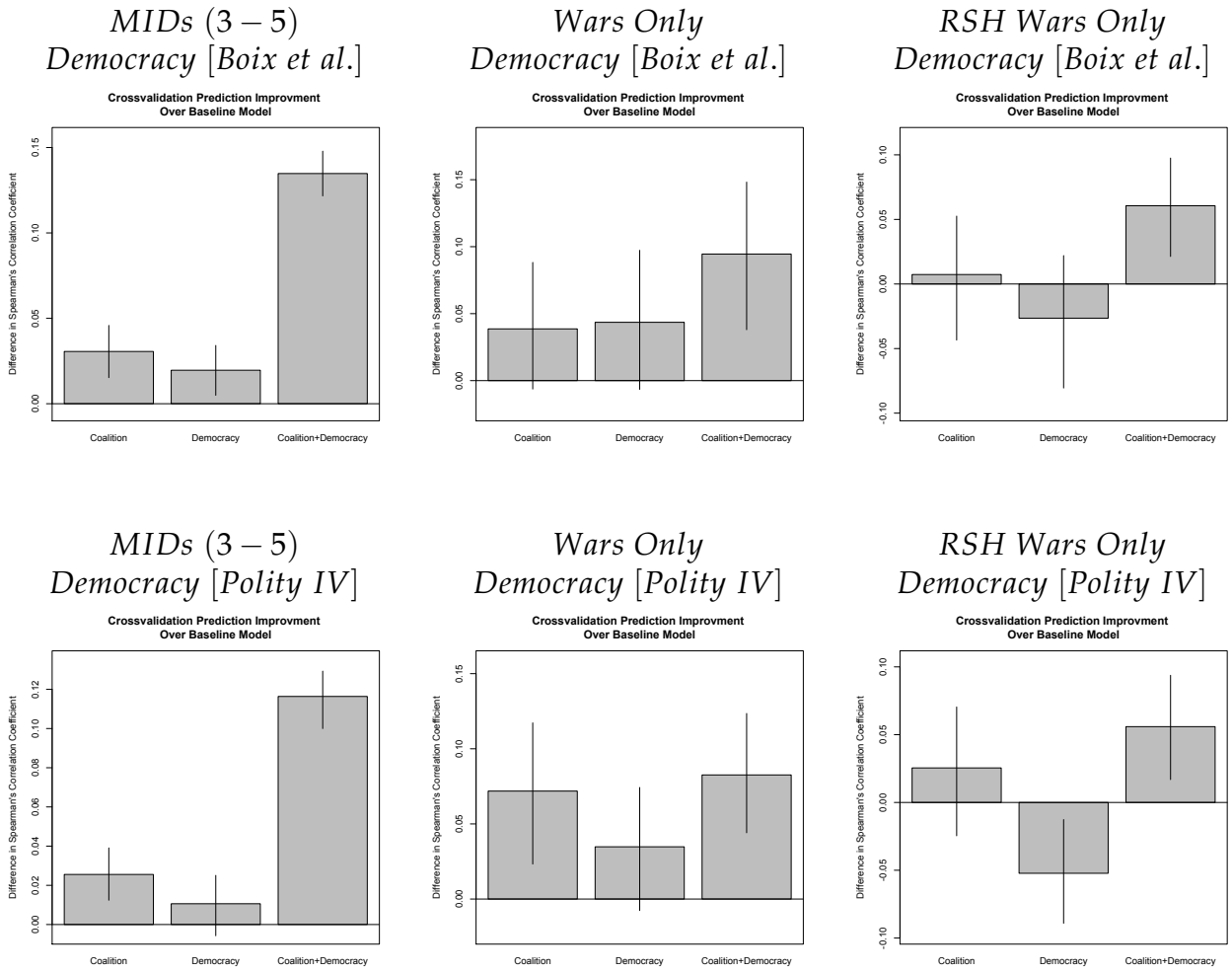

Figure 12: Difference of Spearman's ρ coefficients for three models compared to a baseline model (controls only). The ρ coefficients are correlations between observed and predicted ordered dependent variables. The larger the difference, the greater the improvement in the predictive power of the alternative model relative to the baseline. The MIDS 3-5 and Wars only data are from Maoz (2005) and the RSH Wars are from (Reiter, Stam and Horowitz, 2014).

10.11 Predicted Number of Coalition by CINC Quantile Values

In some models, the military capabilities of the state in question, *CINC score*, has a statistically significant relationship with the number of coalition partners that the state attracts. We display these model relationships graphically here. What we see is that the size and direction of this relationship varies greatly depending on how coalition size is operationalized and which data set is used.

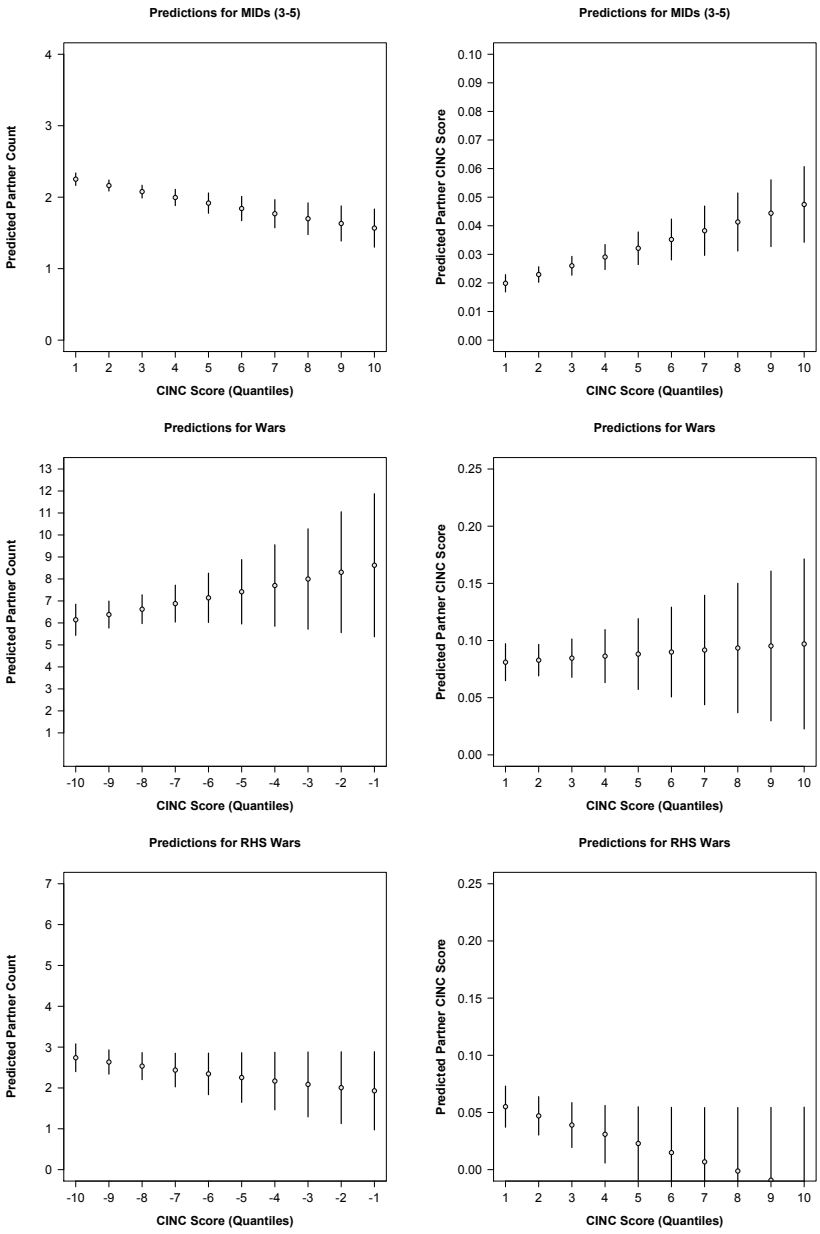

Figure 13: Predicted number of coalition partners (left column) and CINC score for all coalitions partners (right column) as a function of CINC values for MIDs (3-5), Wars only, and RSH Wars. All other variables from the models are held at their mean values.

10.12 Omitting the Control for Dyad MID Propensity

We initially added the control for Dyad MID propensity in response to a reviewer suggestion. However, in the interest of fullest possible transparency, here we also present results without that control. We see results that are substantively the same as those in the body of the paper. To limit the number of tables slightly we display only the results from the bivariate probit models – these results are substantively the same as results from modeling each stage singly.

Table 25: Joint Probability of Partners and Victory: Omitting MID-Propensity Control

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coalition				
Democracy [Polity IV]	0.049***		0.029***	
	(0.017)		(0.0076)	
Democracy [Boix et al.]		0.73***		0.45***
		(0.26)		(0.11)
CINC score	-0.30	-0.35	-0.29	-0.19
	(1.48)	(1.42)	(0.62)	(0.61)
Opponent(s)' CINC score	1.78	1.71	2.03***	2.05***
	(1.09)	(1.09)	(0.69)	(0.69)
Constant	-0.058	-0.29	-1.31***	-1.49***
	(0.22)	(0.23)	(0.10)	(0.092)
DV= Win				
Number of Partners	0.047*	0.043	0.091***	0.091***
	(0.029)	(0.028)	(0.018)	(0.018)
Democracy [Polity IV]	0.0097		-0.00070	
	(0.024)		(0.0057)	
Democracy [Boix et al.]		0.25		-0.075
		(0.38)		(0.088)
CINC score	5.18***	4.91***	3.35***	3.48***
	(1.44)	(1.32)	(0.42)	(0.40)
Opponent(s)' CINC score	-2.60***	-2.47**	-0.061	-0.11
	(0.99)	(0.99)	(0.54)	(0.54)
Troop Quality	0.100	0.087	0.031*	0.025*
	(0.10)	(0.098)	(0.017)	(0.013)
Constant	-0.58**	-0.62***	-1.59***	-1.55***
	(0.25)	(0.22)	(0.055)	(0.055)
Arc-Hyperbolic Tangent	-0.19	-0.16	0.087	0.10
	(0.28)	(0.29)	(0.12)	(0.12)
Observations	344	353	4534	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 26: Joint Probability of Allies and Victory: Omitting MID-Propensity Control

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coaliton (CINC)				
Democracy [Polity IV]	0.11***		0.042***	
	(0.024)		(0.011)	
Democracy [Boix et al.]		1.69***		0.56***
		(0.37)		(0.16)
CINC score	-1.30	-1.54	0.75	0.93
	(2.24)	(2.15)	(0.89)	(0.85)
Opponent(s)' CINC score	0.78	1.01	1.01	1.04
	(1.02)	(1.01)	(0.71)	(0.69)
Constant	-1.47***	-2.10***	-2.17***	-2.36***
	(0.31)	(0.45)	(0.11)	(0.12)
DV = Win				
Partner(s)' CINC score	3.98***	3.87***	4.75***	4.80***
	(1.33)	(1.40)	(0.81)	(0.80)
Democracy [Polity IV]	-0.013		0.0015	
	(0.021)		(0.0056)	
Democracy [Boix et al.]		-0.16		-0.019
		(0.38)		(0.084)
CINC score	5.03***	4.74***	2.37***	2.55***
	(1.33)	(1.24)	(0.54)	(0.51)
Opponent(s)' CINC score	-2.95***	-2.78***	-0.14	-0.19
	(0.83)	(0.80)	(0.56)	(0.55)
Troop Quality	0.10	0.087	0.045***	0.034***
	(0.078)	(0.075)	(0.016)	(0.013)
Constant	-0.62***	-0.53***	-1.55***	-1.53***
	(0.21)	(0.20)	(0.063)	(0.060)
Arc-Hyperbolic Tangent	-0.66***	-0.74**	-0.15	-0.17
	(0.24)	(0.34)	(0.16)	(0.16)
Observations	344	353	4534	4675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Bivariate Probit with errors clustered on dispute-side.

10.13 Splitting the Sample Between Joiners and Initial Participants

We make the theoretical argument in the paper that our argument applies to initiators, targets and joiners. One reader of this work expressed some skepticism that our theory applies to joiners, so we decided to address this empirically as well as theoretically. As in the previous section, we present only the bivariate probit results. The one-stage results are the same. All estimated effects are in the expected direction.

In the sample of joiners only (about half of war participants and 15% of participants in high-level MIDs), we see the effect of coalition size on victory fall from statistical significance ($p<.05$) in the wars-only sample in Table 29, where we operationalize coalition size as the number of powers, but it remains significant even at the 1% level (i.e. $p<.01$) in Table 30, where we operationalize coalition size as partners' summed capabilities. This provides some empirical backing to our theoretical assertion that our theory applies to joiners as well as to initial participants.

Table 27: Joint Probability of Partners and Victory: Joiners Only

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coalition				
Democracy [Polity IV]	0.11**		0.050***	
	(0.046)		(0.012)	
Democracy [Boix et al.]		1.37**		0.85***
		(0.59)		(0.19)
CINC score	-6.95*	-5.80**	-6.63***	-6.79***
	(3.64)	(2.93)	(1.05)	(1.06)
Opponent(s)' CINC score	0.58	0.52	2.21**	2.40**
	(1.54)	(1.62)	(0.99)	(1.00)
Dyad MID Propensity	21.8	22.2	4.05	4.20
	(17.4)	(15.8)	(5.59)	(5.47)
Constant	0.85	0.27	0.31	-0.018
	(0.60)	(0.53)	(0.20)	(0.21)
DV = Win				
Number of Partners	0.040	0.034	0.098***	0.097***
	(0.034)	(0.033)	(0.020)	(0.021)
Democracy [Polity IV]	0.026		0.0037	
	(0.032)		(0.016)	
Democracy [Boix et al.]		0.41		0.044
		(0.46)		(0.24)
CINC score	5.14***	4.76**	2.91***	2.83***
	(1.96)	(1.86)	(1.08)	(1.06)
Opponent(s)' CINC score	-1.43	-1.39	-0.81	-0.74
	(1.41)	(1.37)	(0.88)	(0.91)
Troop Quality	0.078	0.083	0.0051	0.0036
	(0.16)	(0.15)	(0.033)	(0.034)
Dyad MID Propensity	-16.1	-13.7	-3.72	-3.01
	(16.7)	(17.6)	(7.48)	(7.57)
Constant	-0.18	-0.30	-1.21***	-1.22***
	(0.54)	(0.53)	(0.24)	(0.25)
Arc-Hyperbolic Tangent	-0.12	-0.038	-0.14	-0.10
	(0.39)	(0.39)	(0.20)	(0.20)
Observations	170	172	586	598

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 28: Joint Probability of Allies and Victory: Joiners Only

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coalition (CINC)				
Democracy [Polity IV]	0.12***		0.066***	
	(0.032)		(0.020)	
Democracy [Boix et al.]		1.76***		1.01***
		(0.54)		(0.31)
CINC score	-3.09	-3.11*	-2.38	-2.55*
	(2.17)	(1.89)	(1.50)	(1.40)
Opponent(s)' CINC score	-0.48	-0.16	0.43	0.70
	(1.53)	(1.57)	(1.06)	(1.07)
Dyad MID Propensity	11.2	11.0	4.39	3.25
	(13.8)	(13.5)	(9.40)	(9.22)
Constant	-1.26***	-1.90***	-1.45***	-1.78***
	(0.35)	(0.47)	(0.31)	(0.36)
DV = Win				
Partner(s)' CINC score	3.85**	4.26**	2.75***	2.94***
	(1.64)	(1.66)	(0.98)	(0.96)
Democracy [Polity IV]	-0.014		0.014	
	(0.036)		(0.015)	
Democracy [Boix et al.]		-0.33		0.20
		(0.57)		(0.23)
CINC score	5.25***	4.95***	1.04	1.15
	(1.67)	(1.55)	(1.14)	(1.13)
Opponent(s)' CINC score	-1.91	-1.97	-1.20	-1.14
	(1.64)	(1.64)	(1.05)	(1.03)
Troop Quality	0.12	0.10	0.015	0.014
	(0.13)	(0.12)	(0.025)	(0.025)
Dyad MID Propensity	-15.5	-15.0	4.38	4.76
	(16.0)	(16.5)	(7.19)	(7.21)
Constant	-0.34	-0.23	-0.97***	-1.08***
	(0.49)	(0.49)	(0.23)	(0.24)
Arc-Hyperbolic Tangent	-0.63*	-0.83*	-0.16	-0.22
	(0.35)	(0.46)	(0.30)	(0.30)
Observations	170	172	586	598

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Bivariate Probit with errors clustered on dispute-side.

In Tables 29 and 30, we show that results are also robust to limiting the sample to the initial participants (i.e. omitting joiners). Here, our core results are robust across both wars and MIDs, regardless of which way we measure coalition size.

Table 29: Joint Probability of Partners and Victory: Omitting Joiners

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coalition				
Democracy [Polity IV]	0.017		0.027**	
	(0.015)		(0.011)	
Democracy [Boix et al.]		0.22		0.36**
		(0.26)		(0.15)
CINC score	4.55**	4.39**	1.57*	1.68*
	(2.17)	(2.12)	(0.95)	(0.91)
Opponent(s)' CINC score	0.81	0.70	0.64	0.64
	(1.22)	(1.20)	(0.72)	(0.70)
Dyad MID Propensity	24.9***	25.2***	2.12	2.94
	(8.58)	(8.60)	(3.35)	(3.32)
Constant	-1.63***	-1.68***	-1.79***	-1.94***
	(0.34)	(0.33)	(0.19)	(0.15)
DV = Win				
Number of Partners	0.095***	0.089***	0.075***	0.076***
	(0.022)	(0.021)	(0.020)	(0.020)
Democracy [Polity IV]	0.00016		-0.0025	
	(0.018)		(0.0051)	
Democracy [Boix et al.]		0.15		-0.11
		(0.28)		(0.077)
CINC score	4.76	4.77*	3.46***	3.66***
	(2.90)	(2.80)	(0.51)	(0.47)
Opponent(s)' CINC score	-3.79***	-3.56***	0.054	-0.056
	(1.28)	(1.21)	(0.60)	(0.59)
Troop Quality	0.093	0.076	0.041*	0.030**
	(0.085)	(0.082)	(0.022)	(0.015)
Dyad MID Propensity	1.17	1.83	4.81**	4.81**
	(8.67)	(8.61)	(2.06)	(2.00)
Constant	-0.63**	-0.68**	-1.76***	-1.70***
	(0.30)	(0.32)	(0.078)	(0.074)
Arc-Hyperbolic Tangent	-0.68**	-0.64***	0.16	0.16
	(0.26)	(0.24)	(0.13)	(0.13)
Observations	174	181	3948	4077

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 30: Joint Probability of Allies and Victory: Omitting Joiners

	(1) Wars Only	(2) Wars Only	(3) MIDs (3-5)	(4) MIDs (3-5)
DV = Coalition (CINC)				
Democracy [Polity IV]	0.065***		0.021*	
	(0.024)		(0.011)	
Democracy [Boix et al.]		1.49***		0.22
		(0.36)		(0.16)
CINC score	3.31	3.01	2.41***	2.69***
	(3.31)	(4.01)	(0.93)	(0.88)
Opponent(s)' CINC score	0.41	0.94	-0.28	-0.35
	(0.81)	(1.17)	(0.68)	(0.67)
Dyad MID Propensity	-16.0*	-20.7	-4.65	-4.52
	(9.61)	(14.7)	(3.80)	(3.76)
Constant	-1.76***	-2.40***	-2.30***	-2.39***
	(0.37)	(0.50)	(0.17)	(0.15)
DV = Win				
Partner(s)' CINC score	5.02***	2.34	4.94***	4.88***
	(0.64)	(4.30)	(1.20)	(1.21)
Democracy [Polity IV]	-0.0068		-0.00093	
	(0.020)		(0.0057)	
Democracy [Boix et al.]		0.10		-0.058
		(0.38)		(0.086)
CINC score	5.11*	5.73**	2.50***	2.74***
	(2.82)	(2.83)	(0.68)	(0.64)
Opponent(s)' CINC score	-3.74***	-3.30***	0.095	0.0079
	(1.09)	(1.10)	(0.58)	(0.57)
Troop Quality	0.033	0.068	0.052**	0.035**
	(0.060)	(0.12)	(0.021)	(0.016)
Dyad MID Propensity	11.7	9.86	5.66***	5.38***
	(8.27)	(8.53)	(2.13)	(2.05)
Constant	-0.84***	-0.82***	-1.77***	-1.71***
	(0.29)	(0.29)	(0.097)	(0.082)
Arc-Hyperbolic Tangent	-14.5***	-0.25	-0.021	-0.013
	(4.05)	(1.53)	(0.14)	(0.13)
Observations	174	181	3948	4077

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification: Bivariate Probit with errors clustered on dispute-side.

References

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A complete data set of political regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–1554.

Calhoun, C. 1989. "Estimating the Distribution of Desired Family Size and Excess Fertility." *The Journal of Human Resources* 24(4):709–724.

Calhoun, C. A. 1986. *BIVOPROB: Maximum Likelihood Program for Bivariate Ordered- Probit Regression*. Washington D.C.: The Urban Institute.

Choi, Ajin. 2004. "Democratic Synergy and Victory in War, 1816-1992." *International Studies Quarterly* 48:663–682.

Conybeare, John A.C. 1994. "A Portfolio Benefits of Free Riding in Military Alliances." *International Studies Quarterly* 38(3):405–419.

Efron, Bradley. 1983. "Estimating the Error rate of a Prediction Rule: Improvement on Cross-Validation." *Journal of the American Statistical Association* 78(382):316–331.

Geisser, Seymour. 1975. "The Predictive Sample Reuse Method with Applications." *Journal of the American Statistical Association* 70(350):320–328.

Gibler, Douglas and Meredith Reid Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816–2000." *Journal of Peace Research* 41(2):211–222.

Gordon, Phillip I. 2005. *Permutation, Parametric, and Bootstrap Tests of Hypotheses*. Springer-Verlag.

Greene, William H. and David A. Hensher. 2010. *Modeling Ordered Choices: A Primer*. Cambridge, MA: Cambridge.

Hastie, Trevor, Robert J. Tibshirani and Jerome Friedman. 2008. *Elements of Statistical Learning*. New York: Springer-Verlag.

Hill Jr., Daniel W. and Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* .

Hoff, Peter D. and Michael D. Ward. 2004. "Modeling dependencies in international relations networks." *Political Analysis* 12(2):160–175.

Lai, Brian and Dan Reiter. 2000. "Democracy, Political Similarity, and International Alliances, 1816-1992." *Journal of Conflict Resolution* 44(2):203–227.

Maoz, Zeev. 2005. "Dyadic MID Dataset (DYMID) Dataset (version 2.0)." Computer File. `http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html`.

Picard, Richard R. and R. Dennis Cook. 1984. "Cross-Validation of Regression Models." *Journal of The American Statistical Association* 79(387):575–583.

Reiter, Dan, Allan C. Stam and Michael C. Horowitz. 2014. "A Revised Look at Interstate Wars, 1816–2007." *Journal of Conflict Resolution* doi:10.1177/0022002714553107.

Simon, Michael W. and Erik Gartzke. 1996. "Political System Similarity and the Choice of Allies: Do Democracies Flock Together or Do Opposites Attract?" *Journal of Conflict Resolution* 40(4):617–635.

Singer, J. David and Melvin Small. 1966. "Formal Alliances, 1815-1939: A Quantitative Description." *Journal of Peace Research* 3(1):1–31.

Siverson, Randolph M. and Juliann Emmons. 1991. "Birds of a Feather: Democratic Political Systems and Alliance Choices in the Twentieth Century." *Journal of Conflict Resolution* 35(2):285–306.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47(4):1–13.

Ward, Michael D. and Peter D. Hoff. 2007. "Persistent patterns of international commerce." *Journal of Peace Research* 44(2):157–175.

Ward, Michael D., Randolph M. Siverson and Xun Cao. 2007. "Disputes, democracies, and dependencies: A reexamination of the Kantian peace." *American Journal of Political Science* 51(3):583–601.