Appendix

This online appendix contains supplemental and the results of various robustness tests.

Tables Omitted from the Main Text

Tables A1-A4 contain the results from the descriptive analysis discussed within the main text.[1] Tables A5-8 report the coefficient estimates and standard errors pertaining to the models associated with Figure 1 within the primary text.

Table A1.1: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive 1816-2000

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment N = 6,989	54 (0.77)	6,851 (98.03)	6 (0.09)	78 (1.12)
Control N = 38,732	278 (0.72)	38,097 (98.36)	52 (0.13)	305 (0.79)

Binomial (132, 75, 0.5): 0.045
Binomial (583, 305, 0.5): 0.282
Pearson's X^2 : 0.029

Table A1.2: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive, 1816-1945

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment N = 1,061	27 (2.54)	985 (92.84)	2 (0.19)	47 (4.43)
Control N = 15,008	182 (1.21)	14,609 (97.34)	31 (0.21)	186 (1.24)

Binomial (74, 47, 0.5): 0.027
Binomial (368, 186, 0.5): 0.876
Pearson's X^2 : < 0.001

Table A1.3: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive, 1945-2000

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment N = 4,404	19 (0.43)	4,357 (98.93)	4 (0.09)	21 (0.54)
Control N = 23,370	92 (0.39)	23,139 (99.01)	4 (0.09)	118 (0.50)

Binomial (43, 24, 0.5): 0.542
Binomial (210, 118, 0.5): 0.084
Pearson's X^2 : 0.969

Note: Row percentages reported in parentheses.

Table A2.1: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1816-2000

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	22	856	6	47
N = 931	(2.36)	(91.94)	(0.64)	(5.05)
Control	292	7,322	77	303
N = 7,994	(3.65)	(91.59)	(0.96)	(3.79)

Binomial (69, 47, 0.5): 0.004
Binomial (595, 303, 0.5): 0.682
Pearson's X^2 : 0.041

Table A2.2: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1816-1945

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	10	310	4	25
N = 349	(2.87)	(88.83)	(1.15)	(7.16)
Control	180	4,835	40	177
N = 5,232	(3.44)	(92.41)	(0.76)	(3.38)

Binomial (35, 25, 0.5): 0.017
Binomial (357, 177, 0.5): 0.916
Pearson's X^2 : 0.003

Table A2.3: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1946-2000

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	10	303	2	15
N = 330	(3.03)	(91.82)	(0.61)	(4.55)
Control	109	2,452	35	126
N = 2,722	(4.00)	(90.08)	(1.29)	(4.63)

Binomial (25, 24, 0.5): 0.424
Binomial (210, 118, 0.5): 0.297
Pearson's X^2 : 0.585

Note: Row percentages reported in parentheses.

Table A3.1: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1816-2000

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	23	6,918	0	48
N = 6,989	(0.33)	(98.98)	(0.00)	(0.69)
Control	83	38,574	2	73
N = 38,732	(0.21)	(99.59)	(0.01)	(0.19)

Binomial (71, 48, 0.5): 0.004
Binomial (156, 73, 0.5): 0.471
Pearson's X^2 : < 0.001

Table A3.2: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1816-1945

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	14	1,016	0	31
N = 1,061	(1.32)	(95.76)	(0.00)	(2.92)
Control	46	14,907	1	54
N = 15,008	(0.31)	(99.33)	(0.01)	(0.36)

Binomial (45, 31, 0.5): 0.016
Binomial (100, 54, 0.5): 0.484
Pearson's X^2 : < 0.001

Table A3.3: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1945-2000

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	7	4,380	0	17
N = 4,404	(0.16)	(99.46)	(0.00)	(0.39)
Control	36	23,314	1	19
N = 23,370	(0.15)	(99.76)	(<0.01)	(0.08)

Binomial (24, 17, 0.5): 0.064
Binomial (55, 19, 0.5): 0.030
Pearson's X^2 : < 0.001

Note: Row percentages reported in parentheses.

Table A4.1: Longitudinal Effects of Defensive Alliance Formation on War Onset: Conditional Deterrent, 1816-2000

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment N = 931	18 (1.93)	887 (95.27)	2 (0.21)	24 (2.58)
Control N = 7,994	111 (1.39)	7,793 (97.49)	6 (0.08)	84 (1.05)

Binomial (42, 24, 0.5): 0.441
Binomial (195, 84, 0.5): 0.062
Pearson's X^2 : < 0.001

Table A4.2: Longitudinal Effects of Defensive Alliance Formation on War Onset: Conditional Deterrent, 1816-1945

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment N = 349	17 (4.87)	310 (88.83)	2 (0.57)	20 (5.73)
Control N = 5,232	65 (1.24)	5,091 (97.31)	5 (0.10)	71 (1.36)

Binomial (37 20, 0.5): 0.743
Binomial (136, 71, 0.5): 0.668
Pearson's X^2 : < 0.001

Table A4.3: Longitudinal Effects of Defensive Alliance Formation on War Onset: Conditional Deterrent, 1945-2000

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment N = 330	0 (0.00)	328 (99.39)	0 (0.00)	2 (0.61)
Control N = 2,722	46 (1.69)	2,663 (97.83)	1 (0.04)	12 (0.44)

Binomial (2, 2, 0.5): 0.500
Binomial (58, 12, 0.5): <0.001
Pearson's X^2 : 0.114

Note: Row percentages reported in parentheses.

Table A5: Effect of Defensive Alliance Formation on MID Initiation (Figure 1a in main text)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	0.353**	0.248	1.317***	1.034***	-0.398+	-0.270
	(0.126)	(0.153)	(0.168)	(0.208)	(0.227)	(0.278)
Constant	-4.794***	-3.059***	-4.360***	-3.379***	-4.847***	-2.739***
	(0.057)	(0.056)	(0.074)	(0.082)	(0.077)	(0.085)
Observations	44929	8442	15514	4967	25965	2774

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A6: Effect of Defensive Alliance Formation on MID Initiation, Only Alliances with Defensive Terms Included (Figure 1b in main text)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	-0.051	0.151	0.392	0.896***	-0.436*	-0.239
	(0.143)	(0.144)	(0.243)	(0.208)	(0.201)	(0.250)
Constant	-4.838***	-3.054***	-4.228***	-3.461***	-4.848***	-2.633***
	(0.059)	(0.056)	(0.071)	(0.086)	(0.077)	(0.082)
Observations	44433	8606	14782	5020	27742	2762

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A7: Effect of Defensive Alliance Formation on War Onset (Figure 1c in main text)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	1.274***	1.080***	1.986***	1.440***	-0.120	0.891
	(0.203)	(0.240)	(0.252)	(0.261)	(0.625)	(0.813)
Constant	-6.414***	-4.696***	-5.701***	-4.222***	-7.122***	-5.957***
	(0.127)	(0.122)	(0.144)	(0.123)	(0.239)	(0.398)
Observations	44929	8442	15514	4967	25965	2774

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table A8: Effect of Defensive Alliance Formation on War Onset, Only Alliances with Defensive Terms Included (Figure 1d in main text)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	0.502* (0.225)	0.748** (0.232)	0.504 (0.319)	1.234*** (0.254)	-0.405 (0.626)	0.159 (0.768)
Constant	-6.193*** (0.116)	-4.558*** (0.114)	-4.896*** (0.099)	-4.178*** (0.121)	-7.153*** (0.241)	-5.333*** (0.295)
Observations	44433	8606	14782	5020	27742	2762

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Robustness Tests

In this section, we report the results of robustness tests that feature additional models estimated after making modifications to the MID initiation dependent variable in the matching analysis. We first examine whether our results hold among disputes where the initiator actually used force against the target.[2] Put differently, we examine the extent to which alliance formation either deters or fails to deter potential challengers from actually using force against a target. We then determine whether our findings substantively change after including observations of MID initiation where one or both sides were joiners to the original dispute. As in the primary text, we estimate these models using both of our alliance formation variables, one that indicates instances when *any* type of defensive alliance is formed, and the other indicating instances when an alliance with exclusively defensive terms is formed. We also match on the same set of factors, which includes the presence of other alliances; joint democracy; major power status; strategic rivalry; and history of MIDs.

The majority of the relationships reported in the main text hold under each of the modifications outlined above. In our primary analysis, we found that any form of ATOP defensive or conditional deterrent alliance formation was associated with an increase in the likelihood of dispute initiation in the pre-nuclear era (1816-1945). We also found that when

examining only alliances with exclusively defensive terms, conditional deterrence alliances still shared a positive association with dispute initiation, but that ATOP defensive alliances no longer had any effect. Here, we find that each of these results hold, regardless of whether one includes MID joiners or examines only use of force MIDs. This offers additional support for the Steps-to-War Hypothesis.

In our primary analysis, we also found that ATOP defensive alliances reduced the likelihood of dispute initiation in the nuclear era (1946-2000), but that all other alliance types had no effect. Here, this negative relationship becomes statistically insignificant when examining only use of force MIDs without joiners. The support for the Deterrence Hypothesis in the nuclear era therefore appears to be more sensitive to modifications in the dependent variable. Specifically, defensive alliance formation during this time seems to be less likely to deter uses of force than it is lower-level disputes featuring only the threat or display of force.

Finally, when using data from the full sample (1816-2000), we originally found that the formation of any kind of ATOP defensive alliance increased the likelihood of dispute onset, but that exclusively defensive ATOP alliances and all conditional deterrent alliances exerted no significant effect. These results also obtain when examining use of force MIDs and MID joiners, though we do uncover a statistically significant, positive association between conditional deterrent alliances and use of force MIDs when excluding joiners.

In the remainder of this section we report the results of these tests in more detail. We begin with the assessment of MID initiations that feature the use of force and proceed to the analysis including joiners.

Use of Force Disputes

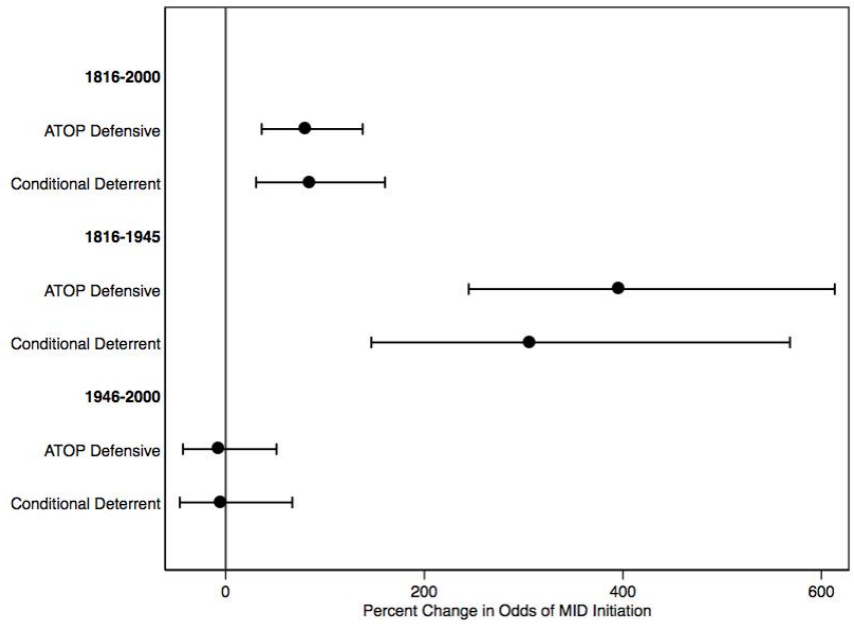
In the main text, our dependent variable records instances where the potential challenger in a particular directed dyad initiated a dispute against a target within the five years after the formation of a defensive alliance. This operationalization of dispute onset includes cases where the challenger state may have only ever threatened or displayed force against the target, without actually resorting to using force. Much of the theoretical literature linking alliances with deterrence, however, is structured around whether a potential challenger actually uses force against the target (e.g. Smith 1995). This also pairs with the terms that comprise defensive alliances, which are often structured around the possibility that a potential challenger attacks an alliance member. To account for this, we also estimated models using a dependent variable that only records whether challengers used force against the target.[3]

Table A9: Effect of Defensive Alliance Formation on Use of Force MID Initiation

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	0.588***	0.613***	1.601***	1.402***	-0.075	-0.054
	(0.142)	(0.176)	(0.186)	(0.254)	(0.250)	(0.290)
Constant	-5.225***	-3.676***	-4.795***	-4.138***	-5.317***	-3.034***
	(0.070)	(0.074)	(0.092)	(0.118)	(0.097)	(0.096)
Observations	45002	8490	15629	4996	25980	2806

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure A1: Effect of Defensive Alliance Formation on Use of Force MID Initiation



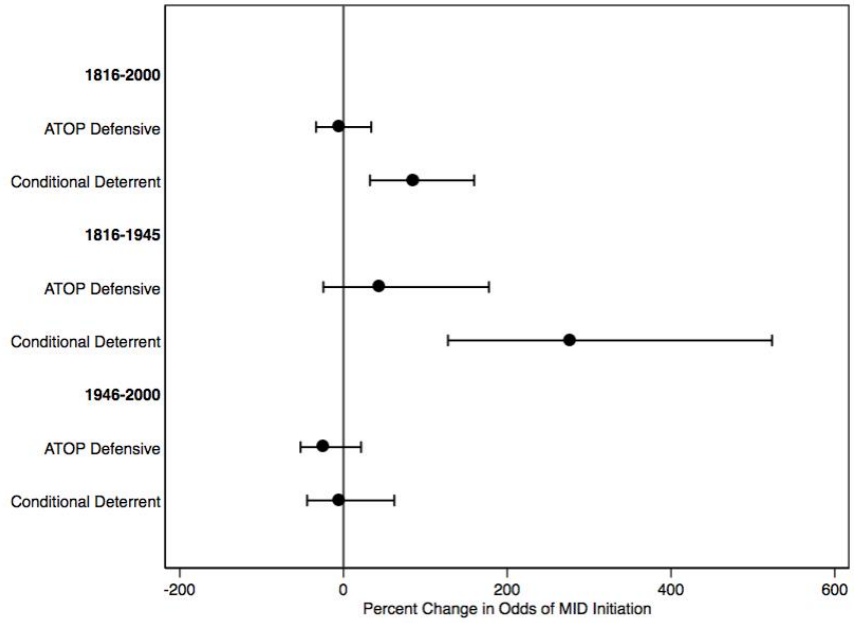
The coefficient estimates pertaining to these models are reported in Tables A9 and A10, while the substantive effect of alliance formation is reported graphically in Figures A1 and A2. Table A9 and Figure A1 report the results for cases where any form of defensive alliance is formed while Table A10 and Figure A2 report the results for purely defensive alliances. Each set of results is similar to those reported in the primary text.

Table A10: Effect of Defensive Alliance Formation on Use of Force MID Initiation, Only Alliances with Defensive Terms Included

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	-0.058	0.617***	0.369	1.326***	-0.274	-0.052
	(0.179)	(0.172)	(0.333)	(0.257)	(0.239)	(0.273)
Constant	-5.291***	-3.825***	-4.863***	-4.275***	-5.335***	-3.008***
	(0.074)	(0.080)	(0.097)	(0.126)	(0.098)	(0.096)
Observations	44484	8655	14793	5049	27753	2794

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure A2: Effect of Defensive Alliance Formation on Use of Force MID Initiation, Only Alliances with Exclusively Defensive Terms Included



When examining all defensive alliance formations during the full time period, ATOP defensive alliances are again positively associated with dispute onset. In contrast to the findings in the text, however, we now find that Benson's conditional deterrent alliances now also share a positive and statistically significant relationship. When examining exclusively defensive alliances, we again find that ATOP defensive alliances have no effect, though conditional deterrent alliances now have a positive association.

In the pre-nuclear era, our substantive findings change little after modifying the dependent variable to include only use of force MIDs. As before, forming a defensive alliance of any kind is positively associated with dispute initiation. When examining only exclusively defensive alliances, conditional deterrent alliances are again positively associated with dispute onset, while ATOP defensive alliances continue to have no effect.

Finally, in the nuclear era, we find that employing use of force MIDs as the dependent variable weakens the support previously uncovered in favor of the deterrence hypothesis. In the main text, the formation of purely defensive ATOP alliances was found to significantly reduce the likelihood of dispute initiation. Here, however, we find that examining use of force MIDs renders this relationship insignificant. This suggests that defensive alliances may be relatively less effective at deterring high-hostility disputes.

In short, limiting the sample of MID initiations to cases involving the use of force by the challenger has, to some extent strengthened the positive association between alliance formation and dispute initiation in the pre-nuclear era, and weakened the negative association thereafter. This offers additional support for the Steps-to-War Hypothesis and weakens that for the Deterrence Hypothesis.

Including Joiners

So far, each of our models has excluded joiners when recording instances of MID initiation. This practice is common in existing research (e.g. Johnson and Leeds 2011, Benson 2012) and is predicated on the notion that the processes behind initiating and joining a dispute are inherently different. As we noted with respect to war, however, deterrence success or failure might also hinge on whether states are likely to join disputes against states that hold defensive alliances. Arguably, the potential problems caused by omitting joiners are more pernicious when studying wars than they are when studying lower-level disputes since states may join a MID before hostilities escalate to the level of interstate war. Even so, we also examine the extent to which our findings change after including joiners in the analysis of MID initiation.

Table A11 and Figure A3 report the results when the treatment indicator pertains to any form of defensive alliance formation with MID initiation as the dependent variable. The

inclusion of joiners substantively altered our results in two ways. First, conditional deterrent alliances are now positively associated with MID initiation for the full time period. Second, ATOP defensive alliances now have a negative association MID initiation in the nuclear era, but this result is only statistically significant at the $p<0.10$ level. Otherwise, the results parallel those reported in the primary text, with ATOP defensive alliances and conditional deterrent alliances each having positive associations with MID initiation in the pre-nuclear era and conditional deterrent alliances have no effect in the nuclear era.

Table A11: Effect of Defensive Alliance Formation on MID Initiation (joiners included)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	0.433*** (0.112)	0.302* (0.137)	1.477*** (0.143)	0.858*** (0.192)	-0.402+ (0.213)	-0.022 (0.240)
Constant	-4.608*** (0.052)	-2.842*** (0.051)	-4.121*** (0.066)	-3.033*** (0.070)	-4.714*** (0.072)	-2.631*** (0.081)
Observations	44929	8442	15514	4967	25965	2774

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure A3: Effect of Defensive Alliance Formation on MID Initiation (with Joiners)

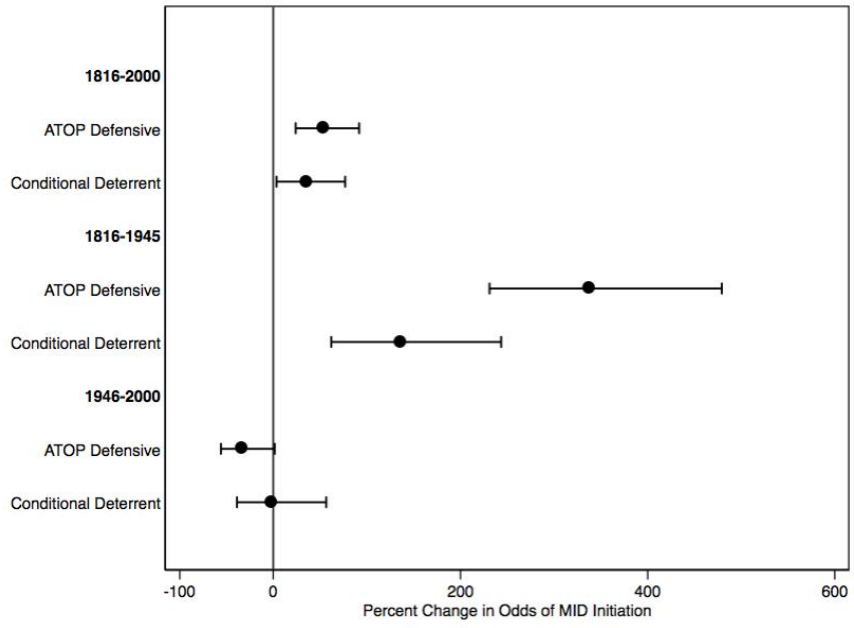


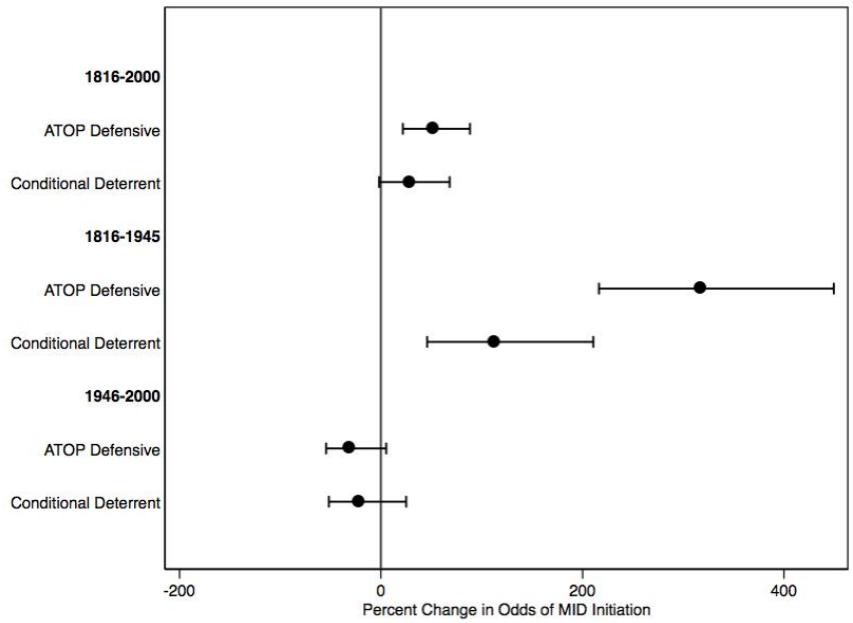
Table A12 and Figure A4 present the results after subsetting cases of MID initiation to only those involving the use of force by the challenger state. In contrast to the use of force analysis in the previous section that excluded joiners, the positive association between conditional deterrent alliances and MID initiation in the full time period is now only statistically significant at the $p<0.10$ level, as is a negative association between ATOP defensive alliances and MID initiation in the nuclear era. Consistent with the analysis excluding joiners, ATOP defensive alliances are still positively associated with MID initiation during the full time period, and that both types of alliances share a positive association in the pre-nuclear era.

Table A12: Effect of Defensive Alliance Formation on MID Initiation (joiners included), Only Alliances with Defensive Terms Included

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	-0.077 (0.129)	0.205 (0.128)	0.270 (0.215)	0.693*** (0.191)	-0.331+ (0.180)	0.084 (0.212)
Constant	-4.614*** (0.053)	-2.829*** (0.050)	-3.867*** (0.060)	-3.091*** (0.072)	-4.709*** (0.072)	-2.542*** (0.078)
Observations	44433	8606	14782	5020	27742	2762

Standard errors in parentheses
+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Figure A4: Effect of Defensive Alliance Formation on Use of Force MID Initiation (with Joiners)



As before, we also report the results including joiners after adjusting the treatment indicator to instances where a defensive alliance with exclusively defensive terms was introduced to a directed dyad that previously had no such alliances. We report these results when using any type of MID initiation (joiners included) as the dependent variable in Table A13 and Figure A5. In contrast to our original findings, ATOP alliances are no longer found to significantly reduce the likelihood of conflict initiation in the nuclear era. Consistent with our primary analysis, however, we find that conditional deterrent alliances significantly increase the likelihood of MID initiation during the pre-nuclear era and that all other relationships are statistically insignificant.

Table A13: Effect of Defensive Alliance Formation on Use of Force MID Initiation (joiners included)

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	0.415***	0.251+	1.428***	0.755***	-0.367+	-0.251
	(0.111)	(0.137)	(0.141)	(0.193)	(0.214)	(0.243)
Constant	-4.582***	-2.809***	-4.061***	-2.969***	-4.749***	-2.460***
	(0.051)	(0.050)	(0.064)	(0.068)	(0.073)	(0.074)
Observations	45002	8490	15629	4996	25980	2806

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure A5: Effect of Defensive Alliance Formation on MID Initiation (with Joiners), Only Alliances with Exclusively Defensive Terms Included

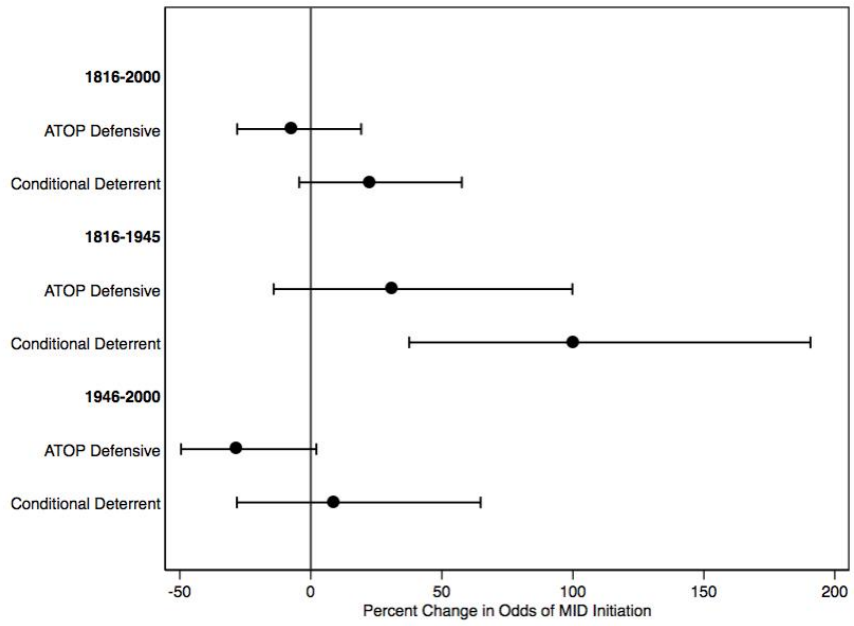


Table A14 and Figure A6 report the results when the treatment indicator pertains to the formation of an alliance with exclusively defensive terms and the dependent variable includes instances of MID joining, but only records cases where the initiator used force against the target. In contrast to the use of force analysis that did not include joiners, we now find that the

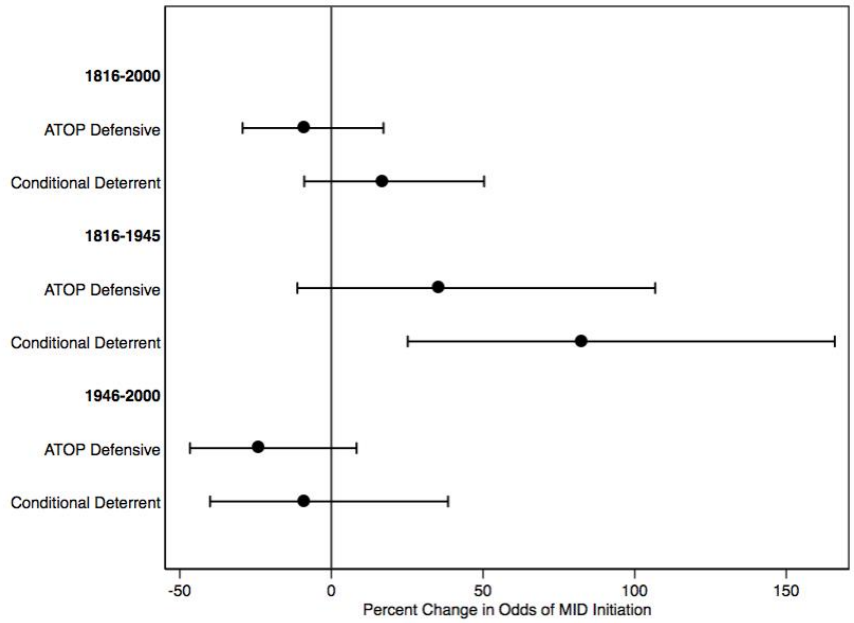
relationship between conditional deterrent alliance formation and dispute onset is statistically significant in the full sample. Consistent with the previous analysis, however, conditional deterrent alliances are positively associated with use of force MID initiation in the pre-nuclear era, but that all other relationships are statistically insignificant.

Table A14: Effect of Defensive Alliance Formation on Use of Force MID Initiation (joiners included), Only Alliances with Defensive Terms Included

	1816-2000		1816-1945		1946-2000	
	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent	ATOP Defensive	Conditional Deterrent
Treatment Indicator	-0.094	0.157	0.304	0.602**	-0.274	-0.093
	(0.129)	(0.128)	(0.216)	(0.192)	(0.181)	(0.213)
Constant	-4.598***	-2.798***	-3.906***	-3.036***	-4.766***	-2.411***
	(0.053)	(0.050)	(0.061)	(0.070)	(0.074)	(0.074)
Observations	44484	8655	14793	5049	27753	2794

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure A6: Effect of Defensive Alliance Formation on Use of Force MID Initiation (with Joiners), Only Alliances with Exclusively Defensive Terms Included



In general, the inclusion of joiners does not greatly modify the findings reported in the main text. In some instances, this modification expands or shifts confidence intervals, which leads to changes in statistical significance. Nevertheless, these changes are relatively small and do not substantially alter the conclusions we draw in the main text.

Supplemental Materials

In addition to the robustness tests conducted above, we provide additional supplemental materials to our primary analysis. These are materials related to our main text that were omitted in the interest of space. We begin with a descriptive analysis of the data pertaining to the formation of alliances with exclusively defensive terms. We then include a description of the variables used in our matching analysis. Finally, we provide an expanded discussion of the manner in which our data were constructed.

Descriptive Analysis – "Pure" Cases of Defensive Alliance Formation

As previously stated, we use two criteria for defining defensive alliance formation, one which required only that a defensive alliance be introduced to a dyad, and the other which required that the formed alliance be "pure" or be devoid of terms which are likely to be considered offensive in nature. In the interest of space, we only reported how alliance formation affected dispute frequency within directed dyads using the former criteria in the main text. Here, we conduct a similar descriptive analysis using the more conservative criteria for defining defensive alliance formation. These results are reported in tables A15-A18. We begin with a discussion with the results pertaining to dispute initiation and proceed to the war analysis.

Table A15.1: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive 1816-2000, Only alliances with exclusively defensive terms included

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	62	7,939	5	56
N = 8,062	(0.77)	(98.47)	(0.06)	(0.69)
Control	322	42,047	56	336
N = 42,761	(0.75)	(98.33)	(0.13)	(0.79)

Binomial (118, 56, 0.5): 0.645
Binomial (658, 336, 0.5): 0.612
Pearson's X^2 : 0.327

Table A15.2: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive, 1816-1945, Only alliances with exclusively defensive terms included

.

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	24	869	1	19
N = 913	(2.63)	(95.18)	(0.11)	(2.08)
Control	191	14,952	33	193
N = 15,369	(1.24)	(97.29)	(0.21)	(1.26)

Binomial (43,19, 0.5): 0.542
Binomial (384,193, 0.5): 0.959
Pearson's X^2 : <0.00

Table A15.3: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: ATOP Defensive, 1945-2000, Only alliances with exclusively defensive terms included

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	30	5,770	4	30
N = 5,834	(0.51)	(98.90)	(0.07)	(0.51)
Control	122	26,499	22	134
N = 26,777	(0.46)	(98.96)	(0.08)	(0.50)

Binomial (60, 30, 0.5): 1.00
Binomial (256,134, 0.5): 0.492
Pearson's X^2 : 0.922

Note: Row percentages reported in parentheses.

Tables A15.1-A15.3 report how the formation of exclusively defensive ATOP alliances affects the frequency of MID initiation in the full time period, pre-nuclear era, and nuclear era respectively. In the full time period, forming a purely defensive alliance has little bearing on MID initiation. While there are more cases experiencing fewer MID initiations after alliance formation, the binomial test indicates that the difference between these two categories is statistically insignificant. Furthermore, there is no significant difference between the distribution of cases in the treatment and control observations, as evidenced by the statistically insignificant value pertaining to the X^2 test.

In the pre-nuclear era, alliance formation again does not lead to a significant increase or reduction in MID initiation, though the distribution between treatment and control groups is now different from one another. This is because treatment observations tend to experience more MID initiations both before *and* after alliance formation. This suggests that treatment observations tend to experience more conflict than control observations. While this might be inferred as weak support for the Steps-to-War Hypothesis, the fact that the same result does not hold in the matching analysis likely suggests it should be rejected in favor of the null hypothesis for this sample of cases.

In the nuclear era, alliance formation does not appear to lead to an increase or reduction in MID initiation, and that the distribution of cases in each category between the treatment and control groups is not statistically significant. This is contrary to the results obtained in the matching analysis, where a weakly significant, negative relationship was found. This indicates that the matching results do not obtain when examining within-case variation of dispute initiation.

Table A16.1: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1816-2000, only alliances with exclusively defensive terms included

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment N = 8,847	28 (2.39)	1,083 (92.41)	7 (0.60)	54 (4.61)
Control N = 1,172	317 (3.58)	8,117 (91.75)	85 (0.96)	328 (3.71)

Binomial (82, 54, 0.5): 0.005
Binomial (645, 328, 0.5): 0.694
Pearson's X^2 : 0.046

Table A16.2: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1816-1945, only alliances with exclusively defensive terms included

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment N = 424	11 (2.59)	384 (90.57)	4 (0.94)	25 (5.90)
Control N = 5,289	181 (3.42)	4,877 (92.21)	42 (0.79)	189 (3.57)

Binomial (36, 25, 0.5): 0.029
Binomial (370, 177, 0.5): 0.436
Pearson's X^2 : 0.083

Table A16.3: Longitudinal Effects of Defensive Alliance Formation on Militarized Dispute Initiation: Conditional Deterrent, 1946-2000 only alliances with exclusively defensive terms included

	Fewer MIDs After	No MID Occurrence Before or After	Same Quantity of MIDs Before and After	More MIDs After
Treatment	15	456	3	22
N = 496	(3.02)	(91.94)	(0.60)	(4.44)
Control	133	3,205	41	139
N = 3,518	(3.78)	(91.10)	(1.17)	(3.95)

Binomial (37, 22, 0.5): 0.324
Binomial (272, 139, 0.5): 0.762
Pearson's X^2 : 0.530

Note: Row percentages reported in parentheses.

Moving to Benson's typology, Tables A16.1-A16.3 reports results pertaining to purely conditional deterrent alliances. In the full time period, the distribution of treatment and control cases across each category is statistically significant and the number of treatment cases featuring and increase in MID initiations after formation is significantly greater than the number of cases featuring a decline. This offers support for the Steps-to-War Hypothesis, but runs contrary to the matching results, which supported the null hypothesis. This latter is perhaps a function of the fact that, while alliance formation appears to lead to a within-case increase in MID initiation, there are also a greater proportion of cases featuring no disputes among the treatment observations than the control observations. Thus, when examining within-case variation exclusively, the balance of evidence is in favor of the Steps-to-War Hypothesis, but after incorporating cross-sectional variation, the evidence is consistent with the null hypothesis. Given the aforementioned advantages associated with the matching approach, we give weight to the results reported in the main text, which leads us to reject the Steps-to-War Hypothesis when examining the 1816-2000 period.

In the pre-nuclear era, the distribution of treatment and control cases is statistically different from one another and that alliance formation leads to a significant increase in MID initiation among the treatment observations. Note also that there are now fewer observations

among the treatment cases that never experience MID initiation. This positive association between alliance formation and MID initiation is consistent with the matching approach and offers support for the Steps-to-War Hypothesis as it pertains to purely conditional deterrent alliances in the pre-nuclear era.

In the nuclear era, we find no clear relationship between the formation of purely conditional deterrent alliances and MID initiation. The formation of an alliance does not lead to a significant increase or reduction in MID initiations among the treatment observations, and there is not a significant difference in the distribution of cases between treatment and control groups. This is again consistent with the matching results, which supported the null hypothesis for this set of cases.

Table A17.1: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1816-2000, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	16	8,016	0	30
N = 8,062	(0.20)	(99.43)	(0.00)	(0.37)
Control	93	42,571	2	95
N = 42,761	(0.22)	(99.56)	(<0.01)	(0.22)

Binomial (46, 30, 0.5): 0.054
Binomial (188, 95, 0.5): 0.942
Pearson's X^2 : 0.082

Table A17.2: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1816-1945, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	6	894	0	13
N = 913	(0.66)	(97.92)	(0.00)	(1.42)
Control	46	15,246	1	76
N = 15,369	(0.30)	(99.20)	(0.01)	(0.49)

Binomial (19,13, 0.5): 0.167
Binomial (122, 76, 0.5): 0.008
Pearson's X^2 : 0.001

Table A17.3: Longitudinal Effects of Defensive Alliance Formation on War Onset:
ATOP Defensive, 1945-2000, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	8	5,809	0	17
N = 5,834	(0.14)	(99.57)	(0.00)	(0.29)
Control	39	26,718	1	19
N = 26,777	(0.15)	(99.78)	(<0.01)	(0.07)

Binomial (25, 17, 0.5): 0.108
Binomial (58, 19, 0.5): 0.012
Pearson's X^2 : < 0.00

Note: Row percentages reported in parentheses.

Proceeding to the analysis of war, Table A17 contains information pertaining to the formation of purely defensive ATOP alliances.[4]. During the full time period, there are nearly twice as many treatment observations featuring more war after formation than before and the difference between these two categories is statistically insignificant. In both the pre-nuclear and nuclear time periods, cases featuring more war after again formation again roughly double those with fewer. While this is consistent with the Steps-to-War Hypothesis, due to a dearth of observations experiencing war, this relationship is statistically insignificant. Nevertheless, it is also worth noting that the proportion of treatment observations in the "more wars after" category is always higher than that of the control observations, suggesting that while alliance formation may not lead to significant within-case increases in war onset, war is still more frequent for directed dyads that recently experienced alliance formation than for those that did not. Overall these results are consistent with those reported in the text: we find support for the Steps-to-War Hypothesis when employing the full time period, but insignificant results in the pre and post nuclear time periods.

Table A18.1: Longitudinal Effects of Defensive Alliance Formation on War Onset:
Conditional Deterrent, 1816-2000, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	17	1,129	1	25
N = 1,172	(1.45)	(96.33)	(0.09)	(2.13)
Control	119	8,630	6	92
N = 8,847	(1.35)	(97.55)	(0.07)	(1.04)

Binomial (42, 25, 0.5): 0.280
Binomial (211, 92, 0.5): 0.073
Pearson's X^2 : 0.012

Table A18.2: Longitudinal Effects of Defensive Alliance Formation on War Onset:
Conditional Deterrent, 1816-1945, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	15	387	1	21
N = 424	(3.54)	(91.27)	(0.24)	(4.95)
Control	70	5,136	5	78
N = 5,289	(1.32)	(97.11)	(0.09)	(1.47)

Binomial (36, 21, 0.5): 0.405
Binomial (148, 78, 0.5): 0.565
Pearson's X^2 : < 0.00

Table A18.3: Longitudinal Effects of Defensive Alliance Formation on War Onset:
Conditional Deterrent, 1945-2000, only alliances with exclusively defensive terms included

	Fewer Wars After	No War Occurrence Before or After	Same Quantity of Wars Before and After	More Wars After
Treatment	1	493	0	2
N = 496	(0.20)	(99.40)	(0.00)	(0.40)
Control	49	3,455	1	13
N = 3,518	(1.39)	(98.21)	(0.03)	(0.37)

Binomial (3, 2, 0.5): 1.00
Binomial (62, 13, 0.5): <0.001
Pearson's X^2 : 0.160

Note: Row percentages reported in parentheses.

The war results in relation to purely conditional deterrent alliances are reported in Table A18. Across each of the three time periods, treatment observations typically experience more wars after alliance formation than before, but the difference between these categories is never statistically significant. Even so, as was the case for the ATOP sample, there is considerable variation between the proportion of treatment and control cases in the "More Wars After" category. In the full time period, 2.13 percent of treatment observations experience more war in the latter-half of the observation period, compared to only 1.04 percent of control observations in the same category. In the pre-nuclear era this trend is even stronger with 4.95 percent of treatment observations featuring more war after formation, compared to only 1.47 percent of

control observations. In the nuclear era, on the other hand, there is little difference between the proportion of cases in the "more wars after" category. Overall, we again find relatively little within-case variation of war onset among the treatment observations, but evidence in the cross-section indicating that alliance formation can be positively linked with war onset in the full and pre-nuclear time periods. This consistent with the matching analysis, which uncovered a positive association between alliance formation and war onset in the full and pre-nuclear samples, but no relationship in the nuclear sample.

Description of Variables

Table A19 contains a list and description of the variables we utilized in the matching analysis.

Table A19: Description of Variables

Variable Name	Description	Details
longitudinal_treatment	Distinguishes "treatment" observations, which experience defensive alliance formation in year 5, from "control" observations, which never experience defensive alliance formation	0: Control 1: Treatment
Descriptive Analysis Variables		
dispute_index	Describes the frequency of MID initiation in the first five years of an observation relative to the last five years. Among treatment observations, this pertains to the five-year periods before and after alliance formation. MIDs are only recorded when the challenger in the directed dyad is the primary challenger in the MID, and the target is the primary target. This information is constructed using the ROLEA and ROLEB variables in the Maoz (2005) dyadic data.	-1: Fewer MID initiations in the first five years than in the last five years 0: No MID initiation occurs 1: The same number of MID initiations in the first and second five year periods 2: More MID initiations in the first five years than in the last five years
warindex	Describes the frequency of war onset in the first five years of an observation relative to the last five years. Wars are recorded regardless of which side in the dyad initiated the MID that escalated to war and whether either side was an originator or joiner. These data are recorded using the HIHOST variable in the Maoz (2005) dyadic data.	-1: Fewer wars in the first five years than in the last five years 0: No war occurs 1: The same number of wars in the first and second five year periods 2: More wars in the first five years than in the last five years

Matching Analysis		
Dependent Variable		
anydispaft	Dichotomous indicator of whether any MID initiations occurred in the last five years of an observation.	0: No MID initiation 1: One or more instances of MID initiation
anydispaft_use	Dichotomous indicator of whether any MID initiations occurred in the last five years of an observation. Only MIDs where the initiator used force against the target are included. This restriction is made using the HIHOST indicator in the Maoz (2005) dyadic data.	0: No use of force MID initiation 1: One or more instances of use of force MID initiation
anywarafter	Dichotomous indicator of whether war occurs in the last five years of an observation.	0: No war 1: At least one war
Variables Matched On		
rivalry_bef	Records how many years in the first half of an observation featured a rivalry between the target and challenger according to the Thompson (2001) rivalry data.	Count of rivalry years, ranging from 0 to 5
disputes_before	Records the number of MID initiations in the first five years of an observation. This is the pre-alliance formation period for treatment observations	Count of MID initiations
jdem*	Dichotomous indicator of whether the challenger and target were both democratic. Democracy is operationalized as a state holding a Polity2 score above 5.	0: Dyad is not a joint democracy 1: Dyad is a joint democracy
ccode1_maj* ccode2_maj	Indicator of Correlates of War major power status. Two variables are constructed, one for the challenger and one for the target.	0: Challenger/target is minor power 1: Challenger/target is major power
pchalneu*	Indicator of whether the challenger holds a neutrality pact. Variable pertains to ATOP alliance typology and is obtained from the Johnson and Leeds (2010) replication data.	0: Absence of challenger neutrality pact. 1: Challenger holds a neutrality pact.
pchaloff*	Indicator of whether an offensive alliance is present. Two variables are constructed, one for the challenger and one for the target. Variable pertains to ATOP alliance typology and is obtained from the Johnson and Leeds (2010) replication data.	0: Absence of challenger/target offensive alliance 1: Challenger/target holds an offensive alliance.
uncondcompel targ_uncondcompel*	Indicator of whether an unconditional compellent alliance is present. Two variables are constructed, one for the challenger and one for the target. Variable pertains to Benson alliance typology and is obtained from the Benson (2011) replication data.	0: Absence of challenger/target unconditional compellent alliance 1: Challenger/target holds an unconditional compellent alliance.
condcompel targ_condcompel*	Indicator of whether an conditional compellent alliance is present. Two variables are constructed, one for the challenger and one for the target. Variable pertains to Benson alliance typology and is obtained from the Benson (2011) replication data.	0: Absence of challenger/target conditional compellent alliance

		1: Challenger/target holds a conditional compellent alliance.
uncondeter*	Indicator of whether the target holds an unconditional deterrent alliance. Variable pertains to Benson alliance typology and is obtained from the Benson (2011) replication data.	0: Absence of unconditional deterrent alliance 1: Target holds an unconditional deterrent alliance.
probabilistic*	Indicator of whether the target holds a probabilistic deterrent alliance. Variable pertains to Benson alliance typology and is obtained from the Benson (2011) replication data.	0: Absence of probabilistic deterrent alliance 1: Target holds a probabilistic deterrent alliance.

- Variable coded in year 5, which marks alliance formation in treatment observations.

Sample Inclusion Criteria

Table A20: Alliance Formation and Inclusion Within the Sample

Directed	Defensive Alliance									
Dyad	*t - 5*	*t - 4*	*t - 3*	*t – 2*	*t – 1*	*t*	*t + 1*	*t + 2*	*t + 3*	*t + 4*
A → B	-	-	-	-	-	X	X	X	X	X
B → A	-	-	-	-	-	X	-	-	X	-
C → D	-	-	-	-	-	X	X	X	X	-
D → C	X	-	-	-	-	X	X	X	X	X

Note: Table displays four hypothetical directed dyads to illustrate which cases would be included within our sample of alliance transitions. Years without alliances are indicated with dashes and years with defensive alliances are indicated with an X. Cases that meet the criteria for being included within the sample have their dyads written in bold text.

As stated in the text, we include observations where a directed dyad transitioned from a period where no alliance was present four of the five years prior to the year in which an alliance is formed, and remained present for at least four of the five years after the alliance was formed. The rule for determining which cases are observed in our sample is represented in Table A20, which reports whether an alliance was present over the course of ten years for four hypothetical directed dyads. Here, the A→B, C→D, and D→C directed dyads would all be observed within the sample. The A→B dyad presents the most straightforward case, where five non-defensive alliance years are followed by five years with a defensive alliance. The C→D and D→C directed

dyads both have anomalous years, i.e. years in the "pre-formation" period with a defensive alliance or years in the "post-formation" period without a defensive alliance. While it is possible for two consecutive years to satisfy these criteria for inclusion- for example, time *t* and *t+1* in the A→B dyad—we avoid duplicates by omitting these overlapping observations. Finally, the B→A dyad would *not* be included in our sample because the defensive alliance does not last for at least four of the five years after formation.

[1] Note that in the tables pertaining to war onset in the nuclear era there is an influx of control observations featuring fewer wars in the first five years of the observation. Of course, we would not expect this result to obtain if the population of control observations were drawn at random. An inspection of the data reveals that the Korean War causes this curious distribution of control observations. Because states like North Korea and China held no defensive alliances and did not exist prior to 1945, all control observations featuring these countries during the time around the Korean War are necessarily classified as having fewer wars in the latter-half of the observation period. If one were to eliminate these observations, the difference in the "more" and "fewer" categories among control observations would no longer be statistically significant, as one would expect.

[2] We do the latter to see whether the results hold if we look only at MIDs with a higher level of hostility, as suggested by one of the reviewers.

[3] This was done using the HiHost indicator associated with the challenger state within the Maoz dyadic data.

[4] As previously discussed, the disproportionally large number of control cases featuring war in the first half of the observation period is due to the Korean War. For this reason, we place less emphasis on the differences in the distribution of treatment and control cases, at least as they pertain to the "fewer wars after" category.