Statistical Appendix for "Deterrence and Restraint: Do Joint Military Exercises Escalate Conflict?"

February 25, 2021

This statistical appendix outlines seven robustness checks to the principal analysis in "Deterrence and Restraint: Do Joint Military Exercises Escalate Conflict?" The key data comes from D'Orazio (2013), based on LexisNexis queries of the AP, AFP, Interfax News, and Xinhua. D'Orazio searched for instances of JMEs by using the following search terms in Lexis-Nexis:

((mil! OR war!) AND (exercis! OR train! OR simulat!) AND NOT (sports OR lifestyle OR tax cuts OR entertainment OR Wall Stree OR baseball)

This process yielded 256,734 documents, which were narrowed down using the support vector machine algorithm, leading to a subset of 12,837 documents that were then hand-coded in order to determine the name, start and end dates, and participants of each exercise. This likely yields a conservative estimate of JMEs, some of which may not be sufficiently newsworthy. Full details on the data collection can be found in D'Orazio (2013: 33-36).

Our key hypotheses are:

Hypothesis 3 (Alliance Reassurance) JMEs in an alliance will reduce the risk that a partner will escalate conflict.

Hypothesis 4 (Alliance Deterrence) JMEs in an alliance will reduce the risk that a target will escalate conflict.

To test them, the main paper regresses *Non-Ally JME*, *Alliance*, and *Ally JME* on *Escalation*, a dichotomous variable measuring whether a state uses military force or declares/joins a war. Following Braithwaite and Lemke (2011), we use a two-stage approach to first control for selection into conflict, and then subsequently assess the effects of our three explanatory variables on escalation.

In addition to providing summary statistics, this appendix presents a number of additional checks:

1. New Operationalization of Explanatory Variables: "Regional" JMEs
2. New Operationalization of Explanatory Variables: Lagged JMEs
3. New Dependent Variable: Combining Reciprocity and Escalation
4. New Model: Multinomial Logistic Regression
5. New Model: Rare Events Logistic Regression
6. Data Generating Process: K-Adic Correction

7. Marginal Effects: System-Wide Restraint on Escalation
8. Robustness Check: Altonji Process for Testing the Sensitivity of Results to Omitted Variables
9. Robustness Check: Molinari Bounds for Testing the Sensitivity of Results to Omitted Variables

For brevity, the robustness checks will present the second-stage results only.

A1 Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Non-Ally JME	0.09	0.28	0	1
Ally JME	0.21	0.41	0	1
Alliances	3.95	5.28	0	48

A2 New Operationalization of Explanatory Variables: "Regional" JMEs

As mentioned in the main text, our data on JMEs does not indicate an exercise's target. To get around this, we assume that JMEs conducted during a conflict are targeted towards the adversary. To determine the sensitivity of our results, we reconstruct the *Non-Ally JME* and *Ally JME* variables, only counting exercises held in the same region as the MID. We then run our first- and second-stage regressions, with the results of the latter presented in Table A1. *Ally JME* retains its negative and significant approach, providing greater confidence that the results in the main text are not sensitive to that initial assumption.

Table A1: Robustness Check: "Regional" JMEs

	Targets	Participants
Non-Ally JME	-0.344	-0.284
	(0.212)	(0.203)
Ally JME	-0.565 ***	-0.366 *
	(0.151)	(0.144)
Ally	-0.006	-0.031 ***
	(0.006)	(0.007)
Joint Democracy	-0.677***	-0.657***
	(0.081)	(0.079)
CINC	12.21 ***	13.39 ***
	(0.976)	(0.937)
UNGA	-0.108 **	-0.111 **
	(0.041)	(0.039)
Trade	0.000002992	0.000002639
	(0.00000372)	(0.000003497)
Lagged DV	6.579 ***	6.164 ***
	(0.084)	(0.084)
Intercept	-7.021 ***	-6.829 ***
	(0.068)	(0.064)
N	542964	542964
AIC	8837.1	9560
$\log L$	-4408.56	-4770.01

Standard errors in parentheses

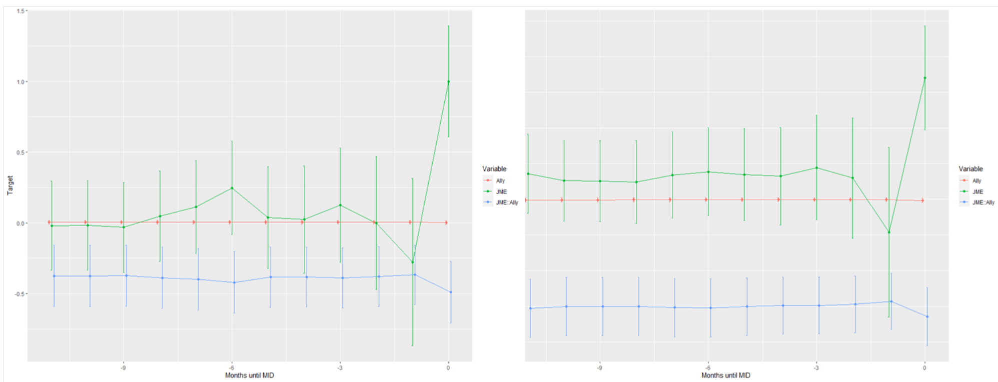
*p<0.1; **p<0.05; ***p<0.01

A3 New Operationalization of Explanatory Variables: Leading JMEs

In Figure A1, we iteratively replace the JME variables with leading versions, ranging from 11 to 0 months before conflict escalation. *JME Ally* possesses consistent effects during this period for both target and participant models, reducing escalation. Further, these graphs cut against the spiral model. That approach would expect the risk of escalation to rise over time. The point estimates for *JME Ally* do do not show significant movement, except perhaps downward during the month of conflict escalation. The confidence intervals also overlap throughout this period.

Non-Ally JME is generally insignificant, except during the month of escalation, where it has a positive effect. Finally, *Alliance*, as expected, consistently has no significant effect on *Escalation*, affirming that military pacts often require additional signaling mechanisms to respond to shorter-term, "local" challenges.

Figure A1: Effects of Leading Versions of *JME* on *Escalation*.



A4 New Dependent Variable: Combining Reciprocity and Escalation

The main paper's models use one operationalization of conflict escalation assessed by Braithwaite and Lemke (2011). To ensure that the results are robust to different operationalizations of the dependent variable, we follow another of Braithewaite and Lemke's examples by combining escalatory responses (the previous DV) with "reciprocal" ones. This occurs when a JME target (Model 1 in Table A2 below) or participant (Model 2) responds to a display of force with their own displays. *Ally JME* remains negative and significant.

Table A2: Robustness Check: New Dependent Variable

	Targets	Participants
Non-Ally JME	-0.060	-0.020
	(0.113)	(0.112)
Ally JME	-0.477***	-0.454***
	(0.089)	(0.088)
Alliances	0.011*	0.023***
	(0.006)	(0.005)
Joint Democracy	-0.630***	-0.758***
	(0.069)	(0.070)
CINC	13.431***	12.461***
	(0.767)	(0.759)
UNGA	0.119***	0.090***
	(0.034)	(0.034)
Trade	0.00001***	0.00001***
	(0.00000)	(0.00000)
Lagged DV	6.218***	6.017***
	(0.085)	(0.085)
Constant	-6.776***	-6.667***
	(0.210)	(0.206)
N	541,920	541,920
AIC	12,262.760	12,502.980
Log Likelihood	-6,091.380	-6,211.488

Standard errors in parentheses

*p<0.1; **p<0.05; ***p<0.01

A5 New Model: Multinomial Logistic Regression

We may be failing to leverage information within our dataset by separating escalation as a distinct category from reciprocity and de-escalation (MID Hostility Levels 1 and 2). Do the effects of our main explanatory variables change once we include these other values on the dependent variable? We use a multinomial logit model to account for this, treating "de-escalation" as the baseline value. As seen in Table A3, *Ally JME* retains its negative and significant effect on *Escalation* in both target and participant models.

Table A3: Robustness Check: Multinomial Logistic Regression

	Targets	Participants
Non-Ally JME	0.014***	0.062***
	(0.00004)	(0.0001)
Ally JME	-0.233***	-0.497***
	(0.0002)	(0.000219)
Alliances	-0.022***	-0.015**
	(0.006)	(0.007)
Joint Democracy	-0.127***	-0.374***
	(0.0001)	(0.0001)
CINC	-2.534***	-2.336***
	(0.00003)	(0.00003)
UNGA	-0.210***	-0.237***
	(0.001)	(0.001)
Trade	-0.000	-0.00000
	(0.00000)	(0.00000)
Lagged DV	2.147***	1.115***
	(0.0001)	(0.0002)
Intercept	1.121***	1.365***
	(0.0003)	(0.0004)
N	541920	541920
AIC	18655.83	18,792.910

Standard errors in parentheses

*p<0.1; **p<0.05; ***p<0.01

A6 New Model: Rare Events Logistic Regression

Conflict escalation occurs in only 0.15 percent of our dataset's observations. A standard logit model could overinflate the significance and substantive effects of our main explanatory variables, and possibly induce an incorrect sign. We use a rare events logit model to correct for this. Table A4 again shows *Ally JME* with a negative and significant association with *Escalation*. We therefore have greater confidence that this relationship is not driven by skewed distributions in the dependent variable.

Table A4: Robustness Check: Rare Events Logistic Regression

	Targets	Participants
Non-Ally JME	-0.045	-0.143
	(0.273)	(0.141)
Ally JME	-0.440*	-0.674*
	(0.117)	(0.169)
Alliances	0.016*	-0.003
	(0.007)	(0.008)
Joint Democracy	-0.742*	-0.717*
	(0.092)	(0.089)
CINC	8.900*	10.59*
	(1.114)	(1.063)
UNGA	-0.050	-0.041
	(0.045)	(0.044)
Trade	0.000	0.000
	(0.000)	(0.000)
Lagged DV	6.598*	6.139*
	(0.092)	(0.092)
Constant	-6.947*	-6.930*
	(0.273)	(0.271)
N	541920	541920
AIC	7754	8402.4

Standard errors in parentheses

*p<0.05

A7 Data Generating Process: K-Adic Correction

Interstate conflict is often a multilateral process. Poast (2010) and Fordham and Poast (2016) demonstrate that we cannot recover the causes or effects of such processes by using dyadic/bilateral units of analysis. We therefore construct a new "k-adic" dataset, subjecting it to a standard model. In both these approaches (and their target and exerciser variants), *Escalate* is negative and significant. Table A5 presents the results.

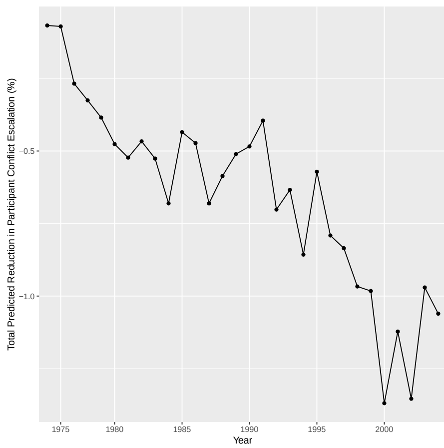
Table A5: Robustness Check: K-Adic Correction for Multilateral Data Generation

	Targets	Participants
Non-Ally JME	-0.295**	-0.384***
	(0.119)	(0.124)
Ally JME	-0.601***	-0.802***
	(0.096)	(0.103)
Alliances	-0.555***	-0.496***
	(0.041)	(0.041)
Joint Democracy	-0.356***	-0.488***
	(0.072)	(0.076)
CINC	6.042***	7.054***
	(1.186)	(1.259)
Rival	-0.130*	-0.145**
	(0.068)	(0.070)
Major Power	0.586***	0.467***
	(0.134)	(0.144)
Contiguity (Land)	4.268***	4.352***
	(0.080)	(0.082)
Contiguity (Sea)	2.467***	2.608***
	(0.123)	(0.126)
Constant	-4.105***	-4.186***
	(0.060)	(0.062)
Observations	33,058	33,058
Log Likelihood	-3,713.566	-3,449.369
Akaike Inf. Crit.	7,447.131	6,918.738

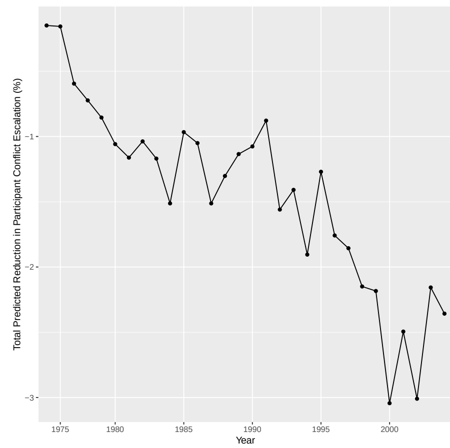
Standard errors in parentheses
*p<0.1; **p<0.05; ***p<0.01

A8 Marginal Effects: System-Wide Restraint on Escalation

The main paper stated that, in 2003, JMEs within an alliance reduced conflict escalation by 3.5% (1% from targets and 2.5% from participants). This calculation was made by multiplying each JME within an alliance by its marginal effect. These conditional marginal effects were used to produce Figure 4 in the main paper. For example, a JME by a state with just one alliance would have around a -.00026% reduction on participant conflict escalation, but one with 20 alliances would have around a -.00024% reduction. Then these values are summed by year, producing Figure A2 below.



(a) Targets



(b) Participants

Figure A2: Annual Reductions in Conflict Escalation due to JMEs within Alliances. The top graph displays the effects of targets, while the bottom graph displays the effects of participants.

A9 Robustness Check: Altonji Process for Testing the Sensitivity of Results to Omitted Variables

The use of observational data raises concerns about selection bias. Without random assignment under experimental conditions, we cannot directly rule out that some omitted variable intermediates JMEs, alliances, and escalation. Altonji, Elder, and Taber (2005) ask: How substantively strong and significant must unobserved factors be to wipe out the effects of our main explanatory variables? If they must be several times stronger, and our controls effectively account for major alternative theoretical explanations, then we can have greater confidence in our explanation.

To formalize their process, Altonji et al. state the following condition:

$$\frac{\mathbb{E}(\epsilon|\text{Predictor} = 1) - \mathbb{E}(\epsilon|\text{Predictor} = 0)}{\text{Var}(\epsilon)} = \frac{\mathbb{E}(X'\gamma|\text{Predictor} = 1) - \mathbb{E}(X'\gamma|\text{Predictor} = 0)}{\text{Var}(X'\gamma)} \tag{1}$$

where X is the matrix of control variables for the outcome equation, γ is a vector of their coefficients, and ϵ is a vector of the residuals from the unobservables. In essence, on the left hand side, we calculate the potential effect that unobserved covariates could have on alliance participation, normalizing that for variation in the error term. On the right hand side, we do the same thing, normalizing for variation in our observed covariates. When this equality holds, a normalized shift in the distribution of unobservables would be equally as powerful as a shift in observables. Altonji et al then transform Equation 1 to ask how large the left hand side must be to explain away our predictor's effects, producing the following ratio, where β is our predictor estimate (say, alliance):

$$\frac{\hat{\beta}}{[\text{Var}(\text{Alliance})/\text{Var}(\text{Residuals})][\mathbb{E}(\epsilon|\text{Alliance} = 1) - \mathbb{E}(\epsilon|\text{Alliance} = 0)]} \tag{2}$$

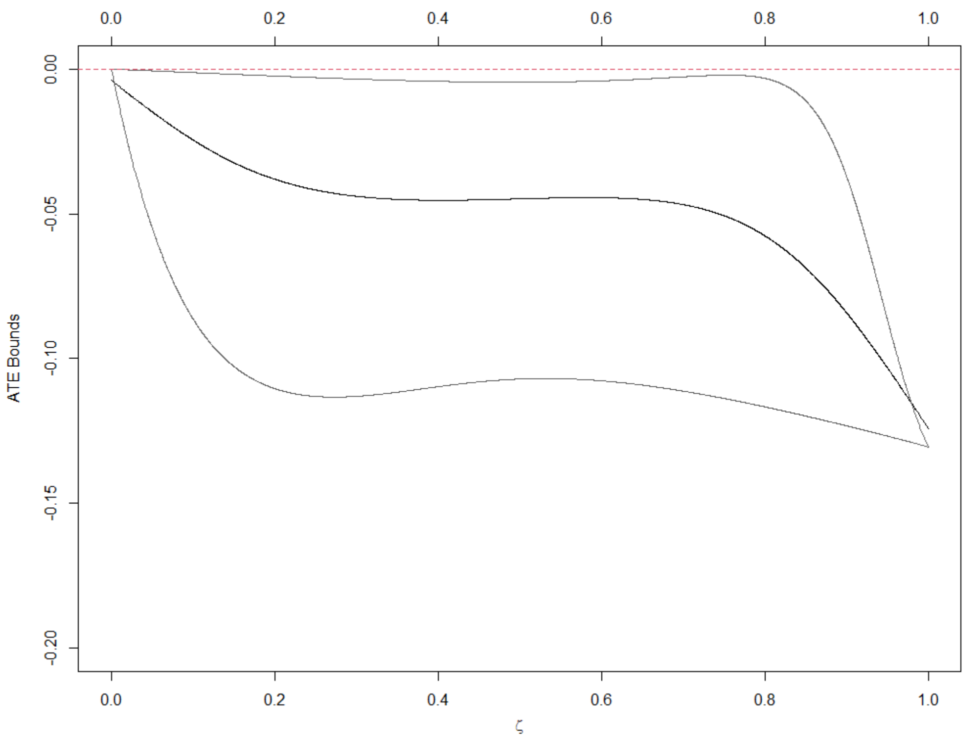
Applying this ratio, we find that (normalized) unobservables must be 23.9 times as strong as *Ally JME* to wipe away its effect on escalation. Given that the model already controls for major alternative accounts, it is unlikely that omitted variables will undermine this result.[1]

[1] For comparison, Altonji, Elder, and Taber (2005) argue that an implied ratio of 3.55 strongly supports their claim.

A10 Robustness Check: Molinari Bounds for Testing the Sensitivity of Results to Omitted Variables

As an additional check on selection bias, we apply modified Molinari bounds developed by Mebane and Poast's (2013). In their paper, the authors directly model the data's sensitivity to violations of the Weak Monotone Treatment Selection (MTS) assumption. Figure A3 visualizes this sensitivity, with the central line demarcating the estimate for the average treatment effect, and the leaf-lines surrounding it 95% posterior intervals. Formally, η is the percentage of the data where selection into treatment does not hold. As it increases, the effect of selection bias attenuates, and Mebane and Poast's (2013) process is agnostic about what causes selection. When selection into a *Ally JME* is complete (i.e. when $\eta = 0$), our effects are indistinguishable from 0. But starting around $\eta = 0.05$ (i.e. 95% of the observations suffer from selection bias), these bounds are statistically distinguishable. In addition, the "flatness" of the curve between $\eta = 0.25$ and 0.75 suggests that our results are insensitive to a range of selection bias effects.

Figure A3: Sensitivity Analysis on Ally JME Using Molinari Bounds.



Supplementary Appendix: References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–184.

Braithwaite, Alex, and Douglas Lemke. 2011. "Unpacking Escalation." *Conflict Management and Peace Science* 28 (2): 111–123.

D'Orazio, Vito. 2013. "International Military Cooperation: From Concepts to Constructs." Ph.D Dissertation, Pennsylvania State University.

Fordham, Benjamin, and Paul Poast. 2016. "All Alliances are Multilateral: Rethinking Alliance Formation." *Journal of Conflict Resolution* 60 (5): 840–865.

Mebane, Walter R., and Paul Poast. 2013. "Causal Inference without Ignorability: Identification with Nonrandom Assignment and Missing Treatment Data." *Political Analysis* 21 (2): 233–251.

Poast, Paul. 2010. "(Mis)Using Dyadic Data to Analyze Multilateral Events." *Political Analysis* 18: 403–425.