Supplemental Material for *Allies as Armaments Explaining the Specialization of State Military Capabilities*

Anonymized

July 18, 2023

This appendix accompanies the paper “Allies as Armaments: Explaining the Specialization of State Military Capabilities”. It provides supplemental information concerning descriptive statistics of the data used in the model and robustness checks and alternate specifications as described in the results section of the manuscript.

Table of contents

1 Descriptive statistics

Table A1: Summary statistics for model variables. Year polynomials are omitted for simplicity.

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Year	45	0	1992.94	12.88	1970.00	1994.00	2014.00
Specialization	5731	10	0.08	0.04	0.00	0.08	0.27
Allies' Mil Spend. (log)	3952	32	11.01	3.09	0.00	11.64	13.97
Allies' CINC Ratio	4866	32	0.92	0.15	0.05	0.98	1.00
GDP (log)	6909	4	24.04	2.16	18.92	23.79	30.42
CINC	7128	1	0.006	0.02	0.000006	0.001	0.23

2 Alternate Model Specifications

We run a set of alternate model specifications as robustness checks.[1] Our results are consistent across alternate modeling specifications including different regression models and control variables. We choose the OLS model specification for the primary results shown in the manuscript given it is the most easily interpretable and consistent with existing research. Explanations for additional model specifications as well as their results are shown below. These alternate model specifications have been chosen based on existing research using similar dependent variables in both ecology and political science (Chiu 2022; Kubinec 2022).

The dependent variable, military specialization, is an entropy-based measures of deviations from a country's count of each military capability from a baseline determined by the composition of their military as well as that of other countries. It is bounded between [0, 1) with 0 representing no specialization (a state's composite military perfectly matches prior expectations) and 1 representing the theoretical entropic maximum. In the observed data (n = 7,203), there are 85 rows where the dependent variable is 0 and none where the dependent variable is 1.

[1]Much of the modeling done here follows the protocol described by Heiss (2021). The author is thankful for their open-source code and strongly suggests readers refer to the original material from which much of this code originates.

2.1 Fractional logit and beta regression

The first set of models provide a robustness check by relaxing the OLS assumption about continuous linearity in the dependent variable. Since the dependent variable is continuous, but bounded between [0, 1) we can use a quasi-binomial fractional logistic regression (Papke and Wooldridge 1996). The coefficients cannot be compared to the original model because they are provided on a log odds scale and can instead be interpreted as percentage change in the dependent variable predicted by a one unit increase in the independent variable value. We can also use a beta regression treating the dependent variable as a proportion bounded between (0, 1) non-inclusive (Grün, Kosmidis, and Zeileis 2012). Since the actual dependent variable does include values of 0, we transform it using the modification suggested by Smithson and Verkuilen (2006). Similar to the fractional logit models, coefficients for the beta regression are on a logit scale. The results from these models (with and without robust standard errors for the fractional logit) are provided in Table A2 and are consistent with the original results provided in the manuscript.[2] Both independent variables retain statistical significance in the expected direction and of comparable magnitudes.

Table A2: Fractional logit and beta regression results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Allies' Mil Spend. (log)	0.019***	0.019***	0.020***			
	(<0.001)	(<0.001)	(<0.001)			
Allies' CINC Ratio				0.243***	0.243***	0.189***
				(<0.001)	(<0.001)	(<0.001)
Democracy	-0.051***	-0.051***	-0.067***	-0.023*	-0.023*	-0.029*

[2]As the beta regression is a distributional, rather than mean-focused regression, it provides a precision parameter that is omitted here (Kneib, Silbersdorff, and Säfken 2021).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	(<0.001)	(<0.001)	(<0.001)	(0.085)	(0.069)	(0.071)
Interstate War (5yr lag)	0.014	0.014	0.001	0.026	0.026	0.015
	(0.679)	(0.724)	(0.982)	(0.452)	(0.521)	(0.722)
GDP (log)	0.157***	0.157***	0.182***	0.169***	0.169***	0.196***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
CINC	0.479*	0.479*	-0.264	0.509*	0.509*	-0.549
	(0.074)	(0.069)	(0.420)	(0.088)	(0.093)	(0.130)
Year	0.070***	0.070***	0.075***	0.073***	0.073***	0.077***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$Year^2$	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$Year^3$	0.00003***	0.00003***	0.00004***	0.00004***	0.00004***	0.00004***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Num.Obs.	3900	3900	3900	3900	3900	3900
Robust SE	No	Yes	No	No	Yes	No
R2			0.290			0.286
AIC			-15843.0			-15807.8
BIC			-15780.3			-15745.1
Log.Lik.		NA			NA	
RMSE	0.03		0.03	0.03		0.03

Note: ^^ * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.2 Zero-inflated and ordered beta regressions

The beta regression cannot handle dependent variables of zero, which is why the above model modifies it. Rather than transforming the zeros, we can also implemented a zero-inflated or ordered beta regression which has the added benefit of allowing us to model a separate data generating process for the country-years with no military specialization at all (Tang et al. 2022; Kubinec 2022). Table A3 shows the results of fully-specified zero-inflated and ordered beta regression models including year fixed effects.[3] As both zero-inflated and ordered beta models are Bayesian, the table specifies only confidence intervals for the main effects since there are no other analogues for frequentist statistics. The results are consistent with those provided in the main text. This suggests that even when relaxing assumptions about the distribution of the dependent variable and the data generating process for states with no specialization, more militarily-capable allies are associated with higher military specialization.

Table A3: Zero-inflated and ordered beta regression results

	Model 1	Model 2	Model 3	Model 4
Intercept	-6.130	-6.319	-6.453	-6.602
	[-6.304, -5.959]	[-6.483, -6.157]	[-6.642, -6.262]	[-6.782, -6.426]
Allies' Mil Spend. (log)	0.017	0.014		
	[0.012, 0.021]	[0.009, 0.019]		
Allies' CINC Ratio			0.202	0.222
			[0.120, 0.287]	[0.131, 0.314]
Democracy	-0.031	-0.054	-0.020	-0.033
	[-0.058, -0.004]	[-0.082, -0.027]	[-0.046, 0.006]	[-0.059, -0.006]

[3] Although year cubic polynomials are used instead of year fixed effects in the main text, they cannot be used here since non-scaled polynomial values complicate model convergence

	Model 1	Model 2	Model 3	Model 4
Interstate War (5yr lag)	-0.004	-0.019	0.005	-0.007
	[-0.090, 0.086]	[-0.093, 0.053]	[-0.079, 0.091]	[-0.081, 0.064]
GDP (log)	0.147	0.156	0.160	0.165
	[0.140, 0.155]	[0.149, 0.164]	[0.153, 0.167]	[0.159, 0.172]
CINC	0.256	0.376	0.469	0.533
	[-0.306, 0.821]	[-0.183, 0.933]	[-0.171, 1.121]	[-0.105, 1.140]
Num.Obs.	3900	3900	3900	3900
R2	0.501	0.531	0.506	0.527
R2 Marg.	0.494	0.524	0.498	0.521
ELPD	8152.8	8128.1	8120.4	8123.5
ELPD s.e.	61.9	60.1	62.1	59.8
LOOIC	-16305.5	-16256.2	-16240.8	-16247.0
LOOIC s.e.	123.9	120.1	124.1	119.5
WAIC	-16305.8	-16256.2	-16241.1	-16247.0
RMSE	0.03	0.03	0.03	0.03

2.3 Fixed effects

The following models serve as robustness checks for temporal and unit-specific trends. While the original models use scaled cubic polynomials to account for temporal trends, the models in Table A4 use year-fixed effects (Carter and Signorino 2010). Furthermore, rather than cluster standard errors at the country-level, models 2 and 4 add robust standard errors. The results are similar to those provided in the main text and consistent with the original hypothesis. I avoid modeling with two-way fixed effects given recent research identifying the shortfalls and biases of such modeling (Callaway and Sant'Anna 2021; Goodman-Bacon 2021; Sun and Abraham 2021; Borusyak, Jaravel, and Spiess 2022). The new estimators that resolve these problems cannot be easily applied here because the treatment variable is continuous and there are more than two time periods (Callaway, Goodman-Bacon, and Sant'Anna 2021).

Table A4: Year-fixed effect regression results

	Model 1	Model 2	Model 3	Model 4
Allies' Mil Spend. (log)	0.001***	0.001***		
	(<0.001)	(<0.001)		
Allies' CINC Ratio			0.025***	0.025***
			(<0.001)	(<0.001)
Democracy	-0.002+	-0.002**	-0.0002	-0.0002
	(0.056)	(0.008)	(0.882)	(0.857)
Interstate War (5yr lag)	0.001	0.001	0.003	0.003
	(0.716)	(0.802)	(0.430)	(0.578)
GDP (log)	0.012***	0.012***	0.013***	0.013***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
CINC	0.183***	0.183***	0.203***	0.203***

	Model 1	Model 2	Model 3	Model 4
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Num.Obs.	3900	3900	3900	3900
R2	0.531		0.529	
R2 Adj.	0.525		0.523	
AIC	-16861.1	-9163.1	-16842.7	-9144.7
BIC	-16823.5	15002.8	-16805.1	15021.3
RMSE	0.03		0.03	

Note: ^^ + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

2.4 Alternate Controls

Table A5: Alternate economic control: GDP per capita

	Model 1	Model 2	Model 3	Model 4
Allies' Mil Spend. (log)	0.004***	0.003**		
	(<0.001)	(0.002)		
Allies' CINC Ratio			−0.045*	−0.007
			(0.022)	(0.589)
Democracy		0.002		0.009*
		(0.599)		(0.024)
Interstate War (5yr lag)		0.007		0.010
		(0.306)		(0.167)
GDP per capita (log)		0.008***		0.011***
		(<0.001)		(<0.001)
CINC		0.728***		0.609**
		(<0.001)		(0.002)
Year		0.005***		0.005***
		(<0.001)		(<0.001)
Year^2		−0.0002***		−0.0002***
		(<0.001)		(<0.001)
Year^3		0.000 003***		0.000 003***
		(<0.001)		(<0.001)
Num.Obs.	4629	3900	4568	3900
AIC	−7400.7	−7922.9	−7079.8	−7765.2
BIC	22 397.6	16 468.8	22 265.1	16 626.4

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

[a] All models include country-clustered standard errors.

Works Cited

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2022. "Revisiting Event Study Designs: Robust and Efficient Estimation." arXiv. https://doi.org/10.48550/arXiv.2108.12419.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment." arXiv. https://doi.org/10.48550/arXiv.2107.02637.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Carter, David B., and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18 (3): 271–92. https://doi.org/10.1093/pan/mpq013.

Chiu, Chun-Huo. 2022. "Incidence-Data-Based Species Richness Estimation via a Beta-Binomial Model." *Methods in Ecology and Evolution* n/a (n/a). https://doi.org/10.1111/2041-210X.13979.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 254–77. https://doi.org/10.1016/j.jeconom.2021.03.014.

Grün, Bettina, Ioannis Kosmidis, and Achim Zeileis. 2012. "Extended Beta Regression in *R* : Shaken, Stirred, Mixed, and Partitioned." *Journal of Statistical Software* 48 (11). https://doi.org/10.18637/jss.v048.i11.

Heiss, Andrew. 2021. "A Guide to Modeling Proportions with Bayesian Beta and Zero-Inflated Beta Regression Models | Andrew Heiss." https://www.andrewheiss.com/blog/2021/11/08/beta-regression-guide/.

Kneib, Thomas, Alexander Silbersdorff, and Benjamin Säfken. 2021. "Rage Against the Mean – A Review of Distributional Regression Approaches." *Econometrics and Statistics*, August.

https://doi.org/10.1016/j.ecosta.2021.07.006.

Kubinec, Robert. 2022. "Ordered Beta Regression: A Parsimonious, Well-Fitting Model for Continuous Data with Lower and Upper Bounds." *Political Analysis*, July, 1–18. https://doi.org/10.1017/pan.2022.20.

Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11 (6): 619–32. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6%3C619::AID-JAE418%3E3.0.CO;2-1.

Smithson, Michael, and Jay Verkuilen. 2006. "A Better Lemon Squeezer? Maximum-likelihood Regression with Beta-Distributed Dependent Variables." *Psychological Methods* 11 (1): 54–71. https://doi.org/10.1037/1082-989X.11.1.54.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 175–99. https://doi.org/10.1016/j.jeconom.2020.09.006.

Tang, Becky, Henry A. Frye, Alan E. Gelfand, and John A. Silander Jr. 2022. "Zero-Inflated Beta Distribution Regression Modeling." arXiv. https://arxiv.org/abs/2112.07249.