Pitfalls of Professionalism?

Military Academies and Coup Risk – Online Appendix

The Online Appendix summarizes several additional analyses and robustness checks, which focus on the unconditional models presented in the main text. The estimations focusing on the conditional effects by regime type are equally robust when implementing any of the following changes in sample or model specifications.

- Appendix Table 1 summarizes the findings from an **out-of-sample 4-fold cross-validation**.
- We re-estimated the unconditional models using a **count variable for military academies** instead of the binary variable (Appendix Table 2).
- Appendix Figure 1 **graphically depicts the curvilinear impact of *Counterbalancing***.
- We also show how the **likelihood of coup attempts shifts over time** (Appendix Figure 2).
- Appendix Table 3 focuses on a **heckman-type probit selection model** on coup attempts and coup outcomes.
- The analyses in Appendix Table 4 **omit the US, France, and Russia**, respectively, from the estimation.
- Another robustness checks summarizes a **random (country-specific) intercept model** (Appendix Table 5).
- The outcome variable is characterized by a **rare-events data generating process**, which we address with an **alternative estimator** in Appendix Table 6.
- We assessed reverse **causality between coups and military academies** using three-stage least-squares regression (Appendix Table 7).

- We **omit states that always/never had a military academy** and estimated a model based on **purely cross-sectional data** only (Appendix Table 8).
- After **including variables capturing internal and external threats**, we assessed the robustness of our main explanatory variable (Appendix Table 9).
- We control for **military regimes** as particularly coup-prone regimes in Appendix Table 10.
- We present a **separation plot** in Appendix Figure 3.
- Finally, Appendix Table 11 re-estimates the unconditional model after employing **genetic one-to-one matching**.

Out-of-Sample 4-Fold Cross-Validation

Hypothesis testing that relies only on statistical significance faces the inherent risk of overfitting to a specific sample's idiosyncrasies and, as argued in Ward, Greenhill, and Bakke (2010), a model may perform not that well with new data (i.e., out-of-sample). To consider out-of-sample heuristics, we also conducted a 4-fold cross-validation quasi-experimental exercise, which we repeated 10 times for the main text's Model 2 and the same model but that omits *Military Academy*.

For this cross-validation, we randomly divide our sample into four segments of about the same size. We then use three random segments out of the four to estimate the model parameters, while the fourth segment is retained for assessing the predictive power of either the main article's Model 2 or that model without *Military Academy* on the pooled subsets. To assess the predictive power out-of-sample, we rely on the area under the Receiver Operator Characteristic (ROC) curve, which ranges from a low value of 0.5 if there is no improvement in predictive power over a random guess to 1.0 for perfect classifications of outcomes. We have repeated this procedure 10 times for each model and also calculated the average values of the AUC measure across these 10 cycle runs. The aim is to assess the predictive power of *Military Academy*: when dropping that item from the main model, the out-of-sample power of the model should decrease. See, for example, Ward, Greenhill, and Bakke (2010: 370) for a more detailed discussion of this approach.

Appendix Table 1 demonstrates that *Military Academy* has predictive out-of-sample power: the average AUC value *decreases* when omitting *Military Academy* from the 4-fold cross validation exercise.

Appendix Table 1. Out-of-Sample 4-Fold Cross Validation

Cycle Run	Main Text Model 2	Model without *Military Academy*
1	0.8249	0.8310
2	0.8331	0.8289
3	0.8308	0.8305
4	0.8318	0.8285
5	0.8361	0.8233
6	0.8328	0.8323
7	0.8328	0.8251
8	0.8325	0.8269
9	0.8297	0.8258
10	0.8265	0.8247
Average Value	0.8311	0.8277

Note: Table entries are area under ROC curve statistics.

The Number of Military Academies

The main explanatory variable in our article is binary – receiving a value of 1 if there is at least one military academy in a given country-year and 0 otherwise. The data from Toronto (2017) also provide information on the *number* of active military academies in a given country-year, though. This count variable ranges in our sample between 0 (such as Japan in 1965 or Switzerland in 2003) and 10 military academies (France as of 2002). The item's median value is 1, e.g., Poland since 1996 or Austria for the entire time period of 1950-2004, with a mean of 1.407 (standard deviation of 1.369).

Appendix Table 2. Count of Military Academies

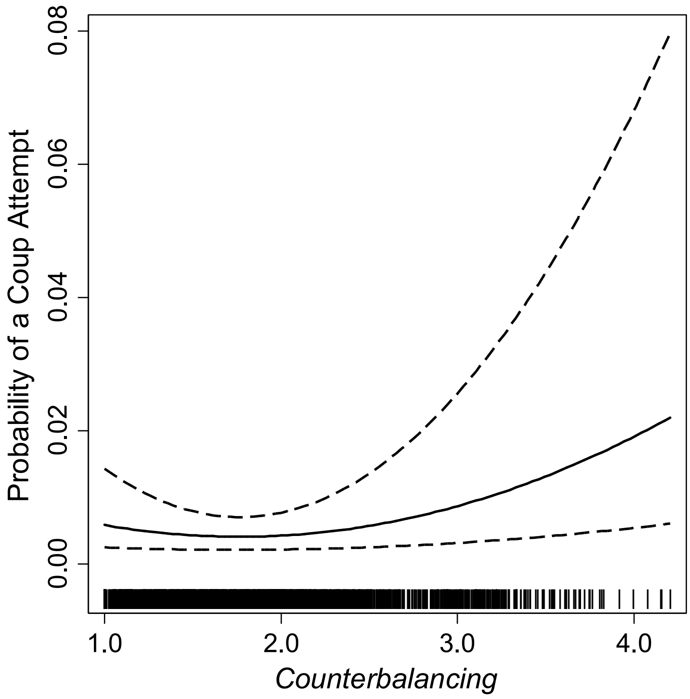
	Model A1	Model A2	Model A3
Military Academy Count	0.115	0.219	0.200
	(0.067)*	(0.073)***	(0.080)**
Change Military Expenditure		-0.133	-0.092
		(0.183)	(0.177)
Soldier Quality		-0.217	-0.397
		(0.071)***	(0.134)***
Military Personnel		-0.204	-0.204
		(0.064)***	(0.079)**
Counterbalancing			-1.395
			(0.837)*
Counterbalancing^2			0.393
			(0.210)*
Change GDP per capita		-3.854	-4.098
		(0.716)***	(0.944)***
GDP per capita		-0.174	-0.113
		(0.109)	(0.127)
Instability		0.124	0.122
		(0.027)***	(0.034)***
Democracy		-0.827	-1.105
		(0.274)***	(0.305)***
Autocracy		-0.113	-0.137
		(0.183)	(0.235)
Constant	-1.841	1.411	3.654
	(0.191)***	(0.794)*	(1.376)***
Obs.	5,088	4,339	3,015
Time Period	1950-2004	1951-2004	1970-2004
Log Pseudo Likelihood	-892.597	-703.025	-385.371
Wald χ^2	102.02	265.30	198.98
Prob>χ^2	0.000	0.000	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Using this count item instead of the binary variable from the main text, we have re-estimated the unconditional models in Appendix Table 2.[1] As demonstrated in Models A1-A3, however, the results based on the count military-academy variable are qualitatively identical to what we discuss in the main text. This applies to both statistical significance and substance. That is, changing *Military Academy Count* from its minimum to its maximum, the simulated risk of a coup is raised by about 15 percentage points (90 percent confidence interval between 0.029 and 0.354) in Model A3.

Appendix Figure 1. The Curvilinear Impact of *Counterbalancing* on Coup Risk



Note: Graph displays predicted probabilities for a coup attempt. Dashed lines pertain to 90 percent confidence interval. The rug plot above the horizontal axis depicts the distribution of *Counterbalancing*.

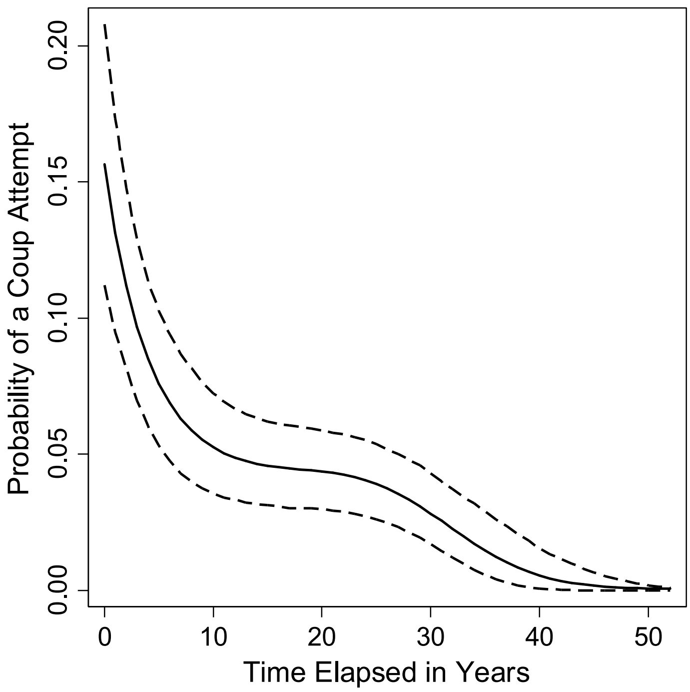
[1] The conditional models are, as with the other analyses in the following, equally robust.

The Curvilinear Impact of *Counterbalancing* and Temporal Dependence of Coups

We have also plotted the curvilinear effect of *Counterbalancing* in Appendix Figure 1, which is based on Model 3 from the main text. Our results show that coup risk first decreases with higher levels of *Counterbalancing*, but then, after a tipping point of about two equally sized military organizations (1.801), increases. Hence, the lowest likelihood of a coup is given in more polarized force-structure environments.

In addition, Appendix Figure 2 shows that the likelihood of a coup d'état rapidly decreases with the time elapsed since the last putsch (if any), but this trend stabilizes in the interval of about 10-20 years, before coup risk slightly decreases again. This graph is based on Model 2 in the main text.

Appendix Figure 2. Temporal Dependencies of Coup Risk



Note: Graph displays predicted probabilities for a coup attempt. Dashed lines pertain to 90 percent confidence interval.

Appendix Table 3. Sample-Selection Model

	Model A4 Selection Stage	Model A4 Outcome Stage
Military Academy	0.345	0.370
	(0.101)***	(0.252)
Change Military Expenditure	-0.070	0.184
	(0.095)	(0.162)
Soldier Quality	-0.104	-0.146
	(0.036)***	(0.097)
Military Personnel	-0.096	-0.005
	(0.028)***	(0.069)
Change GDP per capita	-1.868	-0.056
	(0.338)***	(0.859)
GDP per capita	-0.093	-0.084
	(0.058)	(0.117)
Instability	0.061	-0.015
	(0.013)***	(0.038)
Democracy	-0.411	-0.865
	(0.126)***	(0.292)**
Autocracy	-0.088	0.246
	(0.091)	(0.237)
Constant	0.481	0.883
	(0.422)	(1.078)
Obs.		4,339
Log Pseudo Likelihood		-840.830
Wald χ^2		43.51
Prob>χ^2		0.000

Note: Table entries are coefficients and robust standard errors clustered on country in parentheses. The variables for temporal correction included, but omitted from presentation. The model follows Powell (2012), i.e., we use the temporal controls at the selection stage for identification.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Heckman-Type Selection Model

We also considered both coup attempts and success in one model simultaneously. The factors leading to coups may also influence their outcomes (Powell 2012). Both stages are intertwined and ignoring one of them might bias the results of the other. We use a heckman-type probit setup, which mirrors Powell's (2012) specifications and the exclusion restriction. When running such a model, sample selection is unlikely to strongly affect our results and the two stages of coup attempts and coup outcomes are hardly associated with each other: this is demonstrated by the insignificant estimate for the correlation of the error terms in the two stages (Wald test of independent equations (ρ=0): $\chi^2(1)$=1.25; prob>χ^2= 0.2641). Still, *Military Academy* remains to exert a significant and positive impact on coup attempts.

Appendix Table 4. Omitting Outliers

	Model A5 w/out US	Model A6 w/out France	Model A7 w/out USSR/Russia
Military Academy	0.697	0.697	0.699
	(0.209)***	(0.209)***	(0.209)***
Change Military Expenditure	-0.130	-0.130	-0.127
	(0.181)	(0.181)	(0.182)
Soldier Quality	-0.194	-0.194	-0.189
	(0.074)***	(0.073)***	(0.074)***
Military Personnel	-0.181	-0.180	-0.179
	(0.058)***	(0.057)***	(0.058)***
Change GDP per capita	-3.661	-3.649	-3.680
	(0.667)***	(0.665)***	(0.665)***
GDP per capita	-0.167	-0.164	-0.166
	(0.114)	(0.114)	(0.114)
Instability	0.128	0.128	0.128
	(0.028)***	(0.028)***	(0.028)***
Democracy	-0.823	-0.830	-0.842
	(0.273)***	(0.275)***	(0.275)***
Autocracy	-0.142	-0.147	-0.148
	(0.176)	(0.176)	(0.176)
Constant	0.847	0.826	0.802
	(0.867)	(0.865)	(0.873)
Obs.	4,285	4,285	4,327
Log Pseudo Likelihood	-702.128	-701.697	-701.961
Wald χ^2	296.59	299.62	298.89
Prob>χ^2	0.000	0.000	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Omitting Outliers

We also dropped the US, France, and Russia from a set of models as these countries may appear as outliers. Specifically, the US is coded as 3 to 4 military academies in our sample, France has between 7 and 10 schools, and the USSR/Russia up to three military academies. Yet, neither France nor the US have experienced a coup in the period of our analysis. Including these countries may bias our findings downwards, and removing them entirely from the analysis could lead to a more accurate analysis. However, our main result is not affected in either substance or significance by any these changes as demonstrated in Appendix Table 4. Also, when omitting these countries from the analysis, the unconditional results remain robust.

Country-Specific Random Intercept Model

The number of military academies is presumably country specific and fairly constant, and richer countries are more likely to have such schools. It would follow that something fixed at the level of countries may be explaining coup risk. In order to address this concern, we implemented Model 2 of the main text with a random-effects (random-intercept) approach.[2] To this end, we incorporate a country-specific intercept, which accounts for unobserved heterogeneity at the country level, i.e., certain country-specific factors might influence our results (Rabe-Hesketh and Skrondal 2009). The country-specific random intercept is modeled according to a normal distribution (Gelman and Hill 2009). Our findings pertaining to either the unconditional models (Appendix Table 5) or the conditional ones are identical to those reported in the article, however.

[2] A fixed-effects setup is not possible as more than 50% of the observations are being dropped in such a setup due to the lack of variance for the outcome variable; ultimately, we would induce sample selection bias.

Appendix Table 5. Country-Specific Random Intercept Model

	Model A8
Military Academy	0.642
	(0.245)***
Change Military Expenditure	-0.117
	(0.182)
Soldier Quality	-0.222
	(0.082)***
Military Personnel	-0.227
	(0.073)***
Change GDP per capita	-3.810
	(0.825)***
GDP per capita	-0.284
	(0.133)**
Instability	0.128
	(0.026)***
Democracy	-0.871
	(0.249)***
Autocracy	-0.138
	(0.184)
Constant	1.734
	(1.089)
Obs.	4,339
Log Pseudo Likelihood	-697.393
Country Variance	0.352
Wald χ^2	168.99***
LR Test vs. Logistic Regression	10.11***

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Rare-Events Logistic Regression Model

When discussing our variables in the main text, we also highlighted that coup onset is only observed in about 4.8 percent of all cases in our sample. In other words, the outcome variable is characterized by a rare-events data generating process. As discussed in King and Zeng (2001a, 2001b), this might bias the results obtained via “regular”" estimation procedures like the logistic regression model we employ, as the latter tends to underestimate the probability of onsets. We thus re-estimated the main text's Model 2 with the rare-events logit specification introduced in King and Zeng (2001a, 2001b). The results in Appendix Table 6 are similar to the findings from the article, although we report regular logistic regression models there.

Appendix Table 6. Rare-Events Logistic Regression

	Model A9
Military Academy	0.685
	(0.208)***
Change Military Expenditure	-0.106
	(0.181)
Soldier Quality	-0.193
	(0.073)***
Military Personnel	-0.183
	(0.057)***
Change GDP per capita	-3.661
	(0.666)***
GDP per capita	-0.168
	(0.114)
Instability	0.127
	(0.028)***
Democracy	-0.819
	(0.274)***
Autocracy	-0.140
	(0.176)
Constant	0.880
	(0.862)
Obs.	4,339

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Appendix Table 7. Three-Stage Least-Squares Regression Model

	Model A10 Coup Attempt	Model A10 Military Academy
Coup Attempt		0.956
		(0.111)***
Military Academy	0.053	
	(0.010)***	
Change Military Expenditure		-0.000
		(0.000)
Soldier Quality	-0.017	
	(0.003)***	
Military Personnel	-0.011	
	(0.002)***	
Change GDP per capita	-0.036	
	(0.022)	
GDP per capita		0.048
		(0.011)***
Instability	0.004	
	(0.001)***	
Democracy	-0.036	
	(0.008)***	
Autocracy	-0.005	
	(0.008)	
Constant	0.267	0.466
	(0.027)***	(0.118)***
Obs.		4,339

Note: Table entries are coefficients and standard errors are in parentheses. The variables for temporal correction and fixed effects for the equation on military academies are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Reverse Causality: Three-Stage Least-Squares Regression

We also explored the possibility of reverse causality: it may not (only) be that military academies exert an impact on the likelihood of coups; rather, or additionally, coups influence whether a military academy is established in a country or not. In order to examine this systematically, we estimate a model using three-stage least-squares (3SLS). We considered possible specifications by running multiple models similar to that in the main article. In 3SLS, instruments for endogenous variables are generated by regressing each such variable on all exogenous variables in the system. Here, the endogenous variables are *Military Academy* and *Coup Attempt*. The regression model summarized in Appendix Table 7 is then a re-estimate of Model 2 in the main article using 3SLS. Note that the variables included in the equations must differ in some respects for the model to be identified. Those items included in one, but not the other equation then influence the other equation's outcome indirectly through their dependent variable.

The main finding in Appendix Table 7 is similar to the unconditional result in our article: military academies are positively and significantly related to the risk of a putsch. There are several additional interesting findings as well. On one hand, higher-income countries are positively associated with the establishment of military academies. Recall that this variable was insignificant in the other models with coup onset as the outcome. On the other hand, *Coup Attempt* in the *Military Academy* equation is statistically significant. This supports the view that causality not only flows from *Military Academy* to *Coup Attempt*, but also the other way round.

Omission of Countries Lacking Variation in *Military Academy* and Cross-Sectional Analysis

As discussed in the main text, about half of our sample countries do not display variation on our main explanatory variable: they either never established a military academy or always had one in any year of our sample period. As these countries might bias our results, we have re-estimated the core model while omitting them (Model A11). In addition, Model A12 is based on cross-sectional data only: in order to further address concerns stemming from endogeneity, we have calculated averages for all variables and countries over our observation period and re-estimated the model with OLS. Apart from avoiding some of the methodological challenges associated with time-series cross-sectional data, this approach allows us to deal with outliers in the data: averaging *Military Academy* over entire decades eliminates some of the year-to-year random fluctuations in states' establishment of such organizations. As demonstrated in Appendix Table 8, however, neither change in the model specification affects the substance of our results.

Appendix Table 8. Omitting Countries without Variation and Cross-Sectional Analysis

	Model A11	Model A12
Military Academy	0.615	0.041
	(0.252)**	(0.012)***
Change Military Expenditure	-0.097	-0.001
	(0.204)	(0.000)**
Soldier Quality	-0.258	-0.013
	(0.110)**	(0.007)*
Military Personnel	-0.080	-0.014
	(0.075)	(0.004)***
Change GDP per capita	-3.502	-0.302
	(0.847)***	(0.251)
GDP per capita	-0.272	0.000
	(0.160)*	(0.010)
Instability	0.093	0.014
	(0.036)***	(0.004)***
Democracy	-0.920	-0.007
	(0.344)***	(0.018)
Autocracy	-0.273	0.042
	(0.185)	(0.025)*
Constant	2.128	0.123
	(1.244)*	(0.067)*
Obs.	1,880	104
Log Pseudo Likelihood	-330.423	
Wald χ^2 / F-Test	197.45	8.67
Prob>χ^2 / Prob > F	0.000	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Assessing the Internal and External Threat Environment

Previous research suggests that coup d'états are strongly linked to the internal and external threat environments. Civil conflict or the involvement in an interstate dispute might affect a state's military institutions and how they develop or exert power, and it seems unlikely that military academies are an exception here. In turn, this may well influence the likelihood of a coup attempt via the influence on potential plotters' willingness and opportunity. In this context, note, for example, Svolik (2013) or Acemoglu et al. (2010) who contend that coup risk is higher in times of civil conflict as political leaders cannot credibly commit not to downsize military resources once the threat from rebel groups disappears. And Bell and Sudduth (2017) argue that potential coup plotters have a higher willingness to stage a coup as their benefits derived from the status quo are likely to disappear during ongoing civil war. Similar rationales apply to countries facing external threats (see Pilster and Böhmelt 2012).

We have created two additional controls in light of this discussion. First, using Allansson et al. (2017) and Gleditsch et al. (2002), we consider a binary variable on civil conflict incidence (i.e., onset year and ongoing conflict years are coded as 1). Similarly, using Palmer et al. (2015), we coded a dichotomous item capturing a state's involvement in a militarized interstate dispute in a given country-year. Appendix Table 9 summarizes our findings. On one hand, *Military Academy* remains positively signed and statistically significant. On the other hand, we find evidence for civil conflict being associated with coup attempts (Acemoglu et al. 2010; Svolik 2013; Bell and Sudduth 2017), but the variable linked to interstate disputes is statistically insignificant.

Appendix Table 9. Internal and External Threat Environment

	Model A13
Military Academy	0.852
	(0.210)***
Change Military Expenditure	-0.154
	(0.189)
Soldier Quality	-0.264
	(0.078)***
Military Personnel	-0.237
	(0.067)***
Change GDP per capita	-3.245
	(0.634)***
GDP per capita	-0.044
	(0.121)
Instability	0.097
	(0.031)***
Democracy	-0.785
	(0.267)***
Autocracy	-0.052
	(0.187)
Civil Conflict	0.918
	(0.223)***
Militarized Interstate Dispute	-0.054
	(0.168)
Constant	0.463
	(0.825)
Obs.	4,339
Log Pseudo Likelihood	-689.455
Wald χ^2	355.86
Prob>χ^2	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Appendix Table 10. Military Regimes

	Model A14
Military Academy	0.654
	(0.196)***
Change Military Expenditure	-0.151
	(0.180)
Soldier Quality	-0.263
	(0.076)***
Military Personnel	-0.224
	(0.055)***
Change GDP per capita	-3.917
	(0.744)***
GDP per capita	-0.220
	(0.110)**
Instability	0.132
	(0.027)***
Military Regime	0.419
	(0.184)**
Constant	1.594
	(0.756)**
Obs.	4,445
Log Pseudo Likelihood	-722.790
Wald χ^2	276.06
Prob>χ^2	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

Controlling for Military Regimes

Military regimes are frequently associated with a higher coup risk than other forms of government (Powell 2012: 1028; see also Belkin and Schofer 2003; Thyne 2010). To examine this possibility, we use data from Geddes et al. (2014a, 2014b) who code states as military juntas when "a group of officers decides who will rule and exercise some influence on policies" (Geddes 2003: 51). Instead of the regime-type items (*Democracy* and *Autocracy*), we introduce

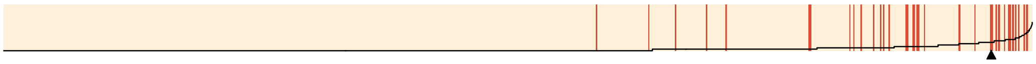
Military Regime into Model A14, using non-military regimes as the baseline category. Appendix Table 10 summarizes the corresponding results. While *Military Academy* is associated with an effect that mirrors our results in the article in both significance and substance, *Military Regime* is, as expected, positively signed and statistically significant. In substantive terms, military regimes are by about 2 percentage points more likely than non-military forms of government to experience a coup, holding all other items constant at their means.

Separation Plot

Separation plots allow the researcher "to evaluate model fit based upon the models' ability to consistently match high-probability predictions to actual occurrences of the event of interest, and low-probability predictions to non-occurrences of the event of interest" (Greenhill, Ward, and Sacks 2011: 990). They rearrange the data so that predicted values are sorted in ascending order (i.e., increase from left to right). In turn, actual instances of the outcome (in our case, coup attempts) and non-events are compared with these predicted values to assess whether and how they correspond.[3] For a model with reasonable predictive power, we would then observe a "clustering" of most events (darker areas or bars) on the right-hand side of the plot. Appendix Figure 3 demonstrates that Model 3 of the main text fits the data reasonably well: most "events" are clustered on the right-hand side of the graph, although a few outliers (i.e., events on the left-hand side in the plot) do exist.

[3] In the words of Greenhill, Ward, and Sacks (2011: 994), "[t]he key idea is that the model's fit (or predictive power) can now be evaluated by simply gauging the extent to which the actual instances of the event are concentrated at the bottom end of the table (right-hand side of the plot), and the nonevents at the top end of the table (left-hand side of the plot). A model with no predictive power – i.e., one whose outcomes can be approximated by a random coin toss – would generate an even distribution of 0s and 1s along the column on the right-hand side."

Appendix Figure 3. Separation Plot – Based on Main Text's Model 3



Note: Solid line pertains to predicted probabilities after reordering of observations. Triangle marks expected number of events.

Matching

Military academies are unlikely to be randomly established by countries (Toronto 2017). The existence of military academies is likely correlated with several confounders that could have positive or negative impacts on coup attempt, which gives rise to an endogeneity issue. As demonstrated by the Turkish case in the main text, the existence of military academies might have been strongly influenced by militaries having capabilities and motivations to (potentially) challenge the incumbent. We sought to address this issue with the simultaneous-equations model above and the analysis of the lag structure of *Military Academy* in the article. Another approach for dealing with non-random assignments is matching. It is "a methodology for reducing bias due to observed covariates in observational studies for causal effects" (Rubin and Thomas 1996: 249). Matching pre-processes the data to form quasi-experimental contrasts by sampling a subset of comparable cases from the overall pool of observations. The observations contained in this subset "match up" each other as closely as possible, i.e., the differences due to other influences are potentially lowered to a minimum. The only exception is that these "most similar" cases differ in whether they received the "treatment:" the establishment of a military academy.

Against this background, the inferences drawn from a matched data set are arguably more informative as differences between the two sets (treatment vs. control group) are solely due to the treatment. We employ genetic one-to-one matching with replacement (Sekhon 2007; Diamond and Sekhon 2013). We used all explanatory variables from the main text's Model 2 to match observations from the treatment group with those from the control group. Ultimately, this approach mirrors the general genetic algorithm used by Sekhon (2007:12ff), which maximizes the smallest p-value for t-tests in each iteration of the matching procedure and, thus, make best use of the balance between treatment and control groups.

Appendix Table 11. Matching

	Model A15
Military Academy	0.746
	(0.290)***
Change Military Expenditure	-0.001
	(0.001)
Soldier Quality	-0.087
	(0.133)
Military Personnel	-0.195
	(0.092)**
Change GDP per capita	-5.357
	(1.166)***
GDP per capita	-0.437
	(0.161)**
Instability	0.135
	(0.037)***
Democracy	-0.631
	(0.450)
Autocracy	0.143
	(0.282)
Constant	1.929
	(1.353)
Obs.	4,700
Log Pseudo Likelihood	-471.183
Wald χ^2	226.90
Prob>χ^2	0.000

Note: Table entries are coefficients. Robust standard errors clustered on country are in parentheses. The variables for temporal correction are included, but omitted from the presentation.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent (two-tailed)

After we have conducted the matching, most explanatory variables' distributions no longer differ significantly between the treated and the control group. Some exceptions exist, however, and we include all explanatory variables used in the matching in the model estimation afterwards to address remaining imbalances. Based on the matching, Ho et al. (2007) suggest using the same parametric estimator with the same set of controls for the matched data that one would have employed in the first place. Appendix Table 11 summarizes our findings. Note that several of those items that were statistically significant in previous analyses no longer achieve statistical significance in Model 15. This is expected, since the matching addresses in the first place the imbalances between treated and control groups. When analyzing the matched sample, *Military Academy* remains robust, however. Hence, even in this quasi-experimental setup that comprehensively controls for the non-random assignment of military academies, we can identify evidence for the proclaimed negative externality of a higher coup risk in this unconditional analysis.

References

Acemoglu, Daron, Andrea Vindigni, and Davide Ticchi. 2010. "Persistence of Civil Wars." *Journal of the European Economic Association* 8 (2-3): 664-676.

Allansson, Marie, Erik Melander, and Lotta Themnér. 2017. "Organized Violence, 1989-2016." *Journal of Peace Research* 54(4): 574-587.

Belkin, Aaron and Evan Schofer. 2003. "Toward a Structural Understanding of Coup Risk." *Journal of Conflict Resolution* 47(5): 594-620.

Bell, Curtis, and Jun Koga Sudduth. 2017. "The Causes and Outcomes of Coup during Civil War." *Journal of Conflict Resolution* 61(7): 1432-1455.

Diamond, Alexis, and Jasjeet Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95(3): 932-945

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor, MI: University of Michigan Press.

Geddes, Barbara, Erica Frantz, and Joseph Wright. 2014a. "Autocratic Breakdown and Regime Transitions: A New Data Set." *Perspectives on Politics* 12(2): 313-331.

Geddes, Barbara, Erica Frantz, and Joseph Wright. 2014b. "Military Rule." *Annual Review of Political Science* 17(1): 147-162.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5): 615-637.

Greenhill, Brian, Michael D. Ward, and Audrey Sacks. 2011. "The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models." *American Journal of Political Science* 55(4): 991-1002.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3): 199-236.

King, Gary, and Langche Zeng. 2001a. "Explaining Rare Events in International Relations." *International Organization* 55(3): 693-715.

King, Gary, and Langche Zeng. 2001b. "Logistic Regression in Rare Events Data." *Political Analysis* 9(2): 137-163.

Palmer, Glenn, Vito D'Orazio, Michael Kenwick, and Matthew Lane. 2015. "The MID4 Dataset, 2002-2010: Procedures, Coding Rules and Description." *Conflict Management and Peace Science* 32(2): 222-242.

Pilster, Ulrich, and Tobias Böhmelt. 2012. "Do Democracies Engage Less in Coup-Proofing? On the Relationship between Regime Type and Civil-Military Relations." *Foreign Policy Analysis* 8(4): 355-372.

Powell, Jonathan. 2012. "Determinants of Attempting and Outcome of Coups d'état." *Journal of Conflict Resolution* 56(6): 1017-1040.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2009. *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press.

Rubin, Donald, and Neal Thomas. 1996. "Matching Using Estimated Propensity Scores: Relating Theory to Practice." *Biometrics* 52(1): 149-264.

Sekhon, Jasjeet. 2007. “Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R.” *Journal of Statistical Software* 42(7): 1-52.

Singh, Naunihal. 2014. *Seizing Power: The Strategic Logic of Military Coups*. Baltimore, MD: Johns Hopkins University Press.

Stepan, Alfred. 1971. *The Military in Politics: Changing Patterns in Brazil*. Princeton, NJ: Princeton University Press.

Svolik, Milan. 2013. “Moral Hazard in Authoritarian Repression and Military Intervention in Politics.” *Journal of Conflict Resolution* 57(5): 765-794.

Thyne, Clayton. 2010. “Supporter of Stability or Agent of Agitation? The Effect of US Foreign Policy on Coups in Latin America, 1960-99.” *Journal of Peace Research* 47(4): 449-461.

Toronto, Nathan W. 2017. “Why Professionalize? Economic Modernization and Military Professionalism.” *Foreign Policy Analysis* 13(4): 854-875.

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): 363-375.