Authoritarian Regimes and Civil-Military Relations:

Explaining Counterbalancing in Autocracies – Online Appendix

The appendix provides a set of robustness checks that further support our theory and results. These include:

- Appendix Table 1 assesses the **predictive power of the personalist-regime** variable via an **out-of-sample 4-fold cross-validation exercise**.
- Appendix Table 2 focuses on the claim that it is particularly **Communist regimes** that invest in counterbalancing.
- In Appendix Table 3, we include a **variable based on the Polity IV data** next to the regime-type dummies, and we **exclude cases that score +6 or higher on the *polity2* scale**.
- We additionally control for **regime duration** in Appendix Table 4
- In Appendix Table 5, we re-estimate the main models with **controls for institutional strength** and **coup history**.
- Appendix Table 6 is based on **country fixed effects** instead of regional fixed effects.
- In Appendix Table 7, we replace the main article's **counterbalancing** variable by the **alternative measure from Pilster and Böhmelt (2011, 2012)**.
- We provide a short **assessment of post-treatment bias**.
- We estimated **count models** using the **number of military organizations** as our dependent variable.
- Finally, we consider **year fixed effects**.

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 (Ward, Greenhill, and Bakke, 2010). To consider out-of-sample heuristics, we conducted a 4-fold cross-validation exercise, which we repeated 10 times for the main text's Model 7 and the same model that omits *Personalist Regime*.

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

Cycle Run	Main Text Model 7	Model without *Personalist Regime*
1	0.9690	0.9663
2	0.9682	0.9671
3	0.9669	0.9676
4	0.9679	0.9667
5	0.9687	0.9674
6	0.9680	0.9673
7	0.9681	0.9675
8	0.9697	0.9673
9	0.9695	0.9666
10	0.9670	0.9660
Average Value	0.9683	0.9670

Note: Table entries are area under ROC curve statistics.

For this cross-validation (see Ward, Greenhill, and Bakke, 2010: 370 for a more detailed discussion of this approach), 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 7 or that model without *Personalist Regime* 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

Personalist Regime: when dropping that item from the main model, the out-of-sample power of the model should decrease.

Appendix Table 1 demonstrates that *Personalist Regime* has predictive out-of-sample power: although the change is rather modest in size (0.00132), the average AUC value decreases when omitting *Personalist Regime* from the 4-fold cross validation exercise.

Appendix Table 2. Communist Regimes

	Appendix Model 1
Communist Regime	-0.166
	(0.426)
Ethnic Fractionalization	-0.581
	(0.980)
Religious Fractionalization	0.963
	(1.194)
Oil	-0.322
	(0.401)
Civil Conflict	0.524
	(0.373)
Militarized Interstate Dispute	-0.097
	(0.215)
Counterbalancing Years	-1.672
	(0.154)***
Counterbalancing Years^2	0.097
	(0.014)***
Counterbalancing Years^3	-0.002
	(0.000)***
Constant	1.818
	(0.825)**
Obs.	1,645
Region Fixed Effects	Yes
Wald χ^2	244.89
Prob > χ^2	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on country in parentheses; region fixed effects included, but omitted from presentation.

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

Appendix Table 3. Including the *Polity2 Score* and Omission of Democratic Country-Years

	Appendix Model 2	Appendix Model 3	Appendix Model 4	Appendix Model 5
Single-Party Regime	-0.750		-0.755	
	(0.438)*		(0.449)*	
Military Regime	-1.406		-1.423	
	(0.719)*		(0.717)**	
Monarchy	-0.647		-0.493	
	(0.818)		(0.795)	
Personalist Regime		0.872		0.867
		(0.441)**		(0.448)*
Polity2	-0.041	-0.048		
	(0.036)	(0.035)		
Ethnic Fractionalization	-0.213	-0.363	-0.508	-0.602
	(0.840)	(0.802)	(0.918)	(0.858)
Religious Fractionalization	0.264	0.291	0.655	0.640
	(1.167)	(1.173)	(1.159)	(1.163)
Oil	-0.355	-0.397	-0.417	-0.493
	(0.368)	(0.381)	(0.368)	(0.381)
Civil Conflict	0.447	0.405	0.357	0.295
	(0.411)	(0.381)	(0.431)	(0.400)
Militarized Interstate Dispute	-0.086	-0.074	-0.097	-0.087
	(0.182)	(0.179)	(0.195)	(0.192)
Counterbalancing Years	-1.622	-1.626	-1.752	-1.761
	(0.154)***	(0.152)***	(0.229)***	(0.228)***
Counterbalancing Years^2	0.090	0.090	0.110	0.110
	(0.014)***	(0.014)***	(0.030)***	(0.030)***
Counterbalancing Years^3	-0.001	-0.001	-0.002	-0.002
	(0.000)***	(0.000)***	(0.001)**	(0.001)**
Constant	2.107	1.341	2.055	1.294
	(0.761)***	(0.891)	(0.801)***	(0.967)
Obs.	1,747	1,747	1,656	1,656
Region Fixed Effects	Yes	Yes	Yes	Yes
Wald χ^2	276.35	238.54	231.98	198.85
Prob > χ^2	0.000	0.000	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on country in parentheses; region fixed effects included, but omitted from presentation.

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

Examination of Communist Regimes

Communist regimes are usually a special type of single-party autocracies. Their exceptional status among single-party dictatorships could also affect how they structure their civil-military relations (e.g., Perlmutter and LeoGrande, 1982; Barany, 1991; Herspring, 1999). Linking this directly to our data, there are several states that have party-based paramilitary organizations or military-youth units that are openly linked to the party; and, in fact, most of these countries could be coded as Communist regimes according to the data in Svolik (2012).

For examining whether our results are indeed driven by Communist regimes (vs. the rest), we took the Communist regime classification from Svolik (2012) and re-estimated our main model again. Appendix Table 2 presents our results. First, note that out of the 312 autocracy-years that are coded as Communist regimes in our data, 34 of them are not classified as single-party regimes (18 of these 34 are coded as personalist regimes, 16 are coded as military juntas). While the control variables remain unchanged, *Communist Regime* is associated with a negative, but statistically insignificant estimate. Therefore, the finding we report in the main text on *Single-Party Regime* is not driven by Communist systems.

Inclusion of *Polity2* Item and Exclusion of Potentially Democratic States

The sample we use for the analyses in the main text are based on Geddes, Wright, and Frantz's (2014) as countries have to be identified as "non-democratic" in a given year. However, according to the Polity IV data set (Marshall and Jaggers, 2013), in which authoritarian regimes tend to range in [-6; -10] and regimes scoring -5 to +5 on the *polity2* scale are generally labeled anocracies, the ideal authoritarian types in Geddes, Wright, and Frantz (2014) span across the entire range of the Polity data, including democracies that range in [6; 10]. For example, Haiti is coded as a military regime by Geddes et al. (2014) in 1994, but receives a *polity2* score of +7 in that year. Hence, we implemented two major changes in our research design. First, we also now require for a country-year to be included in our analysis to have a combined polity

score (*polity2*) lower than +6 in that country-year. Second, we employ the Polity IV data set for an ordinal autocracy-democracy score (Marshall and Jaggers, 2013). This variable is based on the combined polity scale (*polity2*), which ranges in [-10; 10] with -10 indicating fully autocratic states and +10 fully democratic ones. By including the autocratic regime type dummies and the *polity2* item, we address issues of measurement error and control for variation within authoritarian ideal types.

Appendix Table 3 summarizes our results when implementing these changes in the research design. Similar to the findings reported in the main text, *Personalist Regime* is most positively associated with counterbalancing. In fact, all results in Appendix Table 3 mirror what we discussed in the article, including *Single-Party Regime*. Omitting cases based on the Polity IV data set or including the *polity2* item directly into our models does not affect the substance of our results.

Considering Regime Duration

We also assessed the possibility that regime duration influences counterbalancing. An argument on this can be derived from the literature on regime and leader survival (e.g., Bienen and van de Walle, 1989; Svolik, 2012). As a robustness check, we thus created a variable counting the years in power of a regime in place, which is based on Geddes, Wright, and Frantz (2014). Appendix Table 4 summarizes our findings from this additional analysis. Although regime tenure is significantly related to the counterbalancing efforts of an autocracy in this table, controlling for this does not alter our main findings.

Controlling for Institutional Strength and Coup History

Talmadge (2016: 120f) argues that it is particularly weakly institutionalized autocracies that invest in counterbalancing and coup-proofing strategies more strongly. In her view, personalist regimes and military dictatorships are especially more likely to be characterized as "institutionally weak." Conversely, "single-party states are institutionalized in ways that should endow them with the same general invulnerability to coups that stable democracies enjoy and therefore the same latitude to adopt conventional war practices" (Talmadge, 2016: 123). While this argumentation echoes our own and, in fact, is in line with our empirical findings, we also thought about controlling for institutional strength more generally than merely by the ideal autocratic-regime types. That is, using the data by Cheibub et al. (2010), we additionally control for the existence of parties within the legislature (*lparty* variable). This item measures on a nominal scale whether no legislature existed or all members of the legislature are nonpartisan, whether a legislature with only members from the regime party exists, or whether there was a legislature with multiple parties. We created three dummy variables using this information, one for each of the categories, and employ the first category (no legislature) as the baseline. This rationale behind this treatment of institutional strength is based on, among others, Talmadge (2016) who sees regimes with a single individual's domination of the state apparatus or with few/no parties as particularly fragile. Therefore, a "real" party system with policy alternatives may offer more stability and strength.

Similarly, following the same rationale, Models 10 and 11 in this appendix control for coup history. Coup-proofing tactics are taken by leaders for the purpose of reducing coup risk, while the literature also points out that coup-proofing might cause counter-coups. Therefore, it seems important to control for a country's coup history in the analyses. To this end, by relying on the coup-attempt data set by Powell and Thyne (2011), we constructed a variable that counts the years since the last coup and include a logged version into our estimation. The rationale for taking the natural log is that there are likely to be decreasing returns with more time elapsed.

Appendix Table 4. Considering Regime Duration

	Appendix Model 6	Appendix Model 7
Single-Party Regime	-1.076	
	(0.446)**	
Military Regime	-0.864	
	(0.640)	
Monarchy	-1.670	
	(0.840)**	
Personalist Regime		1.051
		(0.414)**
Regime Duration	0.798	0.746
	(0.143)***	(0.119)***
Ethnic Fractionalization	0.370	0.272
	(0.908)	(0.920)
Religious Fractionalization	-0.527	-0.167
	(1.145)	(1.090)
Oil	-0.612	-0.516
	(0.363)*	(0.393)
Civil Conflict	0.778	0.805
	(0.362)**	(0.368)**
Militarized Interstate Dispute	-0.096	-0.099
	(0.217)	(0.215)
Counterbalancing Years	-1.589	-1.582
	(0.148)***	(0.147)***
Counterbalancing Years^2	0.085	0.085
	(0.011)***	(0.011)***
Counterbalancing Years^3	-0.001	-0.001
	(0.000)***	(0.000)***
Constant	0.948	-0.178
	(0.692)	(0.766)
Obs.	1,810	1,810
Region Fixed Effects	Yes	Yes
Wald χ^2	314.34	304.73
Prob > χ^2	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on country in parentheses; region fixed effects included, but omitted from presentation.

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

Appendix Table 5. Controlling for Institutional Strength and *Coup History*

	Appendix Model 8	Appendix Model 9	Appendix Model 10	Appendix Model 11
Single-Party Regime	-0.720		-1.348	
	(0.423)*		(0.512)***	
Military Regime	-1.275		-0.786	
	(0.690)*		(0.661)	
Monarchy	-0.374		-1.572	
	(0.778)		(0.850)*	
Personalist Regime		0.805		1.209
		(0.425)*		(0.447)***
Legislature w/ Regime Members	0.603	0.753		
	(0.427)	(0.417)*		
Legislature w/ Multiple Parties	0.553	0.636		
	(0.320)*	(0.339)*		
Coup History			0.865	0.785
			(0.176)***	(0.136)***
Ethnic Fractionalization	-0.075	-0.115	0.220	0.287
	(0.822)	(0.783)	(0.932)	(0.953)
Religious Fractionalization	0.455	0.277	-0.067	0.058
	(1.120)	(1.142)	(1.115)	(1.078)
Oil	-0.343	-0.398	-0.542	-0.489
	(0.354)	(0.360)	(0.334)	(0.361)
Civil Conflict	0.431	0.381	0.649	0.665
	(0.392)	(0.371)	(0.398)*	(0.405)*
Militarized Interstate Dispute	-0.028	-0.016	-0.131	-0.127
	(0.192)	(0.188)	(0.188)	(0.191)
Counterbalancing Years	-1.583	-1.591	-1.548	-1.546
	(0.146)***	(0.146)***	(0.144)***	(0.144)***
Counterbalancing Years^2	0.088	0.088	0.083	0.083
	(0.011)***	(0.011)***	(0.011)***	(0.011)***
Counterbalancing Years^3	-0.001	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Constant	1.480	0.743	0.578	-0.632
	(0.651)**	(0.810)	(0.758)	(0.913)
Obs.	1,760	1,760	1,810	1,810
Region Fixed Effects	Yes	Yes	Yes	Yes
Wald χ^2	376.33	302.75	290.43	277.04
Prob > χ^2	0.000	0.000	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on country in parentheses; region fixed effects included, but omitted from presentation.

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

Appendix Table 5 summarizes the findings from these two robustness checks. While the variables on institutional strength are partly statistically significant, the regime-type items are

unaltered in either the direction of their impact or their substantive effect. To be clear, the personalist regime-type dummy variable and the *lparty* item from Cheibub et al. (2010) are positively correlated, which suggests that a higher level of state weakness is usually associated with these kinds of regimes, but other forces seem to be at work next to this factor, too. Otherwise, the correlation between the two items would be higher, and the statistical and substantive impact of the institution-variables in Appendix Table 5 larger. Also, note that the coup-years item is statistically significant.

Appendix Table 6. Country Fixed Effects

	Appendix Model 12	Appendix Model 13
Single-Party Regime	-0.724	
	(0.376)*	
Military Regime	-0.451	
	(0.602)	
Monarchy	-0.469	
	(0.895)	
Personalist Regime		0.647
		(0.345)*
Civil Conflict	0.646	0.642
	(0.321)**	(0.317)**
Militarized Interstate Dispute	-0.315	-0.312
	(0.211)	(0.211)
Counterbalancing Years	-1.258	-1.257
	(0.104)***	(0.104)***
Counterbalancing Years^2	0.072	0.072
	(0.008)***	(0.008)***
Counterbalancing Years^3	-0.001	-0.001
	(0.000)***	(0.000)***
Obs.	1,023	1,023
Country Fixed Effects	Yes	Yes
LR χ^2	523.33	532.07
Prob > χ^2	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors in parentheses; country fixed effects included, but omitted from presentation.

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

Including Country Fixed Effects

Dummy variables for countries can control for unobserved, time-invariant unit-level effects. However, they are associated with two potential problems in our context. First, adding binary variables for states into a dichotomous dependent-variable model leads to a potentially large drop-out in the number of observations: due to perfect prediction, all states that never change their behavior on the outcome variable over time are omitted from the analysis. This refers to all countries that either always or never have invested in counterbalancing. Eventually, if the decrease in the number of observations is large, this may suggest that there is selection on stable country-level characteristics. Second, country fixed effects would make the estimation of those variables that change over time only slowly inefficient (Plümper and Troeger, 2011): while the autocracy items are time-varying, such changes only occur rarely in the period under study.

That being said, despite the potential problems due to the inclusion of country fixed effects, we performed such an analysis. Note that all time-invariant variables are directly omitted in such models, i.e., *Ethnic Fractionalization*, *Religious Fractionalization*, and *Oil* are not part of the following models. Appendix Table 6 presents the fixed effects estimations. There is a decrease in the sample size, but our core findings remain robust.

Alternative Counterbalancing Measure

We replaced the counterbalancing measures from De Bruin (2018) by the data introduced in Pilster and Böhmelt (2011, 2012) that also measure counterbalancing. The latter measure uses "information on both the number of rivaling military organizations and their respective strengths to capture the degree to which a state divides its military manpower into rivaling organizations" (Pilster and Böhmelt, 2012: 360). More specifically, the level of counterbalancing is based on the effective number of ground-combat compatible military organizations using the following formula:

$$Counterbalancing = \frac{1}{\sum_j s_{jit}^2}$$

where s_{jit} is the personnel share of a ground-combat compatible military or paramilitary organization *j* in country *i* in year *t*. A *Counterbalancing* value of 1 pertains to "only one effective ground-combat military organization, while higher values signify that rivaling military organizations do exist. The higher the value of that measure, the higher the effective number of military organizations in a country. This in turn signifies higher institutional coup-proofing efforts in the form of creating an artificial balance between various rivaling military organizations" (Pilster and Böhmelt, 2012: 361). *Counterbalancing* is an interval variable that can only assume non-negative values, while a value of 0 is possible (no military organization in a country).

Before coming to our re-estimation of our models with this alternative dependent variable, we briefly compare the data from Pilster and Böhmelt (2011, 2012) with De Bruin (2018) and discuss the respective strength and weaknesses. The data from Pilster and Böhmelt (2011, 2012) are based on the International Institute for Strategic Studies' (IISS) Military Balance and capture 199 countries between 1970 and 2018. De Bruin (2018), on the other hand, has compiled data on 264 security forces in 65 randomly selected developing states with populations over 250,000 between 1960 and 2010. In light of this, a first difference between the two data sets is the different spatial and temporal coverage: Pilster and Böhmelt (2011, 2012) do not focus on developing countries only, but all states that are covered by the Military Balance; their data also cover more recent years. Having said that, De Bruin (2018) goes back until 1960, while Pilster and Böhmelt's (2011, 2012) data have coverage until 1970 only.

In addition, De Bruin (2018: 1440) seems to employ a more restrictive definition in order to capture states' counterbalancing efforts than Pilster and Böhmelt (2011, 2012) have. She focuses on "counterweights" that have to meet two criteria to be identified: "(1) the forces are independent from military command. Instead, operational control, which refers to the ability to

initiate and terminate military operations, rests with the executive, interior ministry, or other government body besides the defense ministry, which controls the military. (2) The forces are deployed within 60 miles of the capital. In practice, this excludes border and frontier guard forces in most years, as well as rural militia." Eventually, De Bruin (2018excludes forces controlled by the military and that are outside the capital. Conversely, Pilster and Böhmelt (2011, 2012) identified *all* ground-combat compatible military organizations of a country using the Military Balance, i.e., army and marine corps troops. But they did not consider "coast guards" and any organizations referring to the terms "port," "aviation," "fishery," "maritime," "marine police," "air police," "air wing," or "naval" in their names. For navies including marine units or air forces including paratroopers, they considered only these units to be ground-combat compatible.

The two main advantages of De Bruin's data are that uniform criteria for including forces in the data are employed and that there are details on the command and deployment of each security organization, which allows for distinguishing between units created for counterbalancing and those that are not. Moreover, another asset is "the triangulation of over 1,200 primary and secondary sources, which reduces measurement error" (De Bruin, 2018: 1435). On the other hand, the quality of the IISS's Military Balance source has been thoroughly discussed in the literature (e.g., Pilster and Böhmelt, 2011; De Bruin, 2018; Narang and Talmadge, 2017). Pilster and Böhmelt (2012: 361f) encountered in particular three kinds of problems. First, although the Military Balance restricts the classification of a paramilitary organization to "forces whose training, organization, equipment, and control suggest they may be used to support, or replace, regular military forces," some of the forces listed in this category actually seem to have a focus on primarily civilian tasks. Second, there are various paramilitary organizations that do not display any information on their troop sizes. Third, there is a certain level of inconsistency and fluctuation in the Military Balance's reporting of the troop levels of

both regular and paramilitary forces. Finally, De Bruin (2018: 1435) demonstrates that quite a few of the security forces she has coded are missing in the Military Balance, "while others are included years after they have been disbanded." However, while inaccuracies do exist in these data and some of the organizations established by leaders to coup-proof their military may be omitted, as in any other data source, Pilster and Böhmelt (2012) also show with various analyses that the degree of inconsistency is low. And, in fact, the two data sets correlate positively at r=0.2249 ($p<0.000$) when focusing on the Pilster and Böhmelt (2011, 2012) measure we use below and De Bruin's (2018) we employ in the main text.

While we rely on De Bruin's (2018) alternative data collection efforts in the main text, Appendix Table 7 comprises re-estimations of the models in Appendix Table 6, but the outcome variable is now based on Pilster and Böhmelt (2011, 2012). As these models rely on fixed effects for countries and years and we include a temporally lagged dependent variable, we use the Arellano-Bond estimator that is linear dynamic panel-data model where the unobserved panel-level effects are correlated with the lags of the dependent variable. The country fixed effects follow the specifications in Appendix Table 6, while the year fixed effects replace the counterbalancing cubic polynomials and address any system-wide shocks that affect all states simultaneously in their counterbalancing decisions. The lagged dependent variable addresses country-specific path dependencies in counterbalancing. Next to the two core models in Appendix Table 7 (which mirror the specifications in Appendix Table 6), we also summarize two additional models that further control for GDP per capita and population size (both taken from the World Bank Development Indicators). Note, however, that these two items are likely to be suffering from post-treatment bias, which we discuss in the next section.

Appendix Table 7. Alternative Outcome Variable

	Appendix Model 14	Appendix Model 15	Appendix Model 16	Appendix Model 17

Single-Party Regime	-0.206		-0.181	
	(0.087)**		(0.091)**	
Military Regime	-0.152		-0.165	
	(0.092)*		(0.109)	
Monarchy	0.113		0.047	
	(0.214)		(0.253)	
Personalist Regime		0.169		0.167
		(0.071)**		(0.079)**
Civil Conflict	-0.016	-0.015	-0.023	-0.023
	(0.030)	(0.030)	(0.030)	(0.030)
Militarized Interstate Dispute	-0.001	-0.001	-0.002	-0.003
	(0.018)	(0.018)	(0.019)	(0.019)
GDP per capita (ln)			-0.013	-0.013
			(0.035)	(0.035)
Population (ln)			0.083	0.083
			(0.146)	(0.145)
Lagged Dependent Variable	0.606	0.608	0.582	0.582
	(0.023)***	(0.023)***	(0.024)***	(0.024)***
Constant	0.646	0.506	-0.578	-0.652
	(0.089)***	(0.061)***	(2.345)	(2.334)
Obs.	2,607	2,607	2,119	2,119
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
χ^2	1,212.69	1,208.86	1,055.91	1,055.41
Prob > χ^2	0.000	0.000	0.000	0.000

Note: Table entries are regression coefficients; standard errors in parentheses; country and year fixed effects included, but omitted from presentation.

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

Appendix Table 7 summarizes our findings. Despite the different outcome variable in Models 14-17 of this appendix, the results mirror those reported in the main text: single-party regimes are negatively associated with counterbalancing, while personalist dictators invest more strongly than other autocrats in structural coup-proofing. We conclude that our findings are robust to alternative data for the dependent variable.

Assessment of Post-Treatment Bias

We also assess the possibility of post-treatment bias. Post-treatment occurs when the consequence of treatment is included as a control in a regression model. Post-treatment bias

has been widely discussed in the literature (King and Zeng, 2006; Elwert and Winship, 2014; Acharya, Blackwell, and Sen, 2016): it biases the estimates as a post-treatment control will soak up some of the effect of the treatment. To circumvent this problem, the main analyses only incorporate a series of controls, next to the fixed effects, which are unlikely to be driven by the authoritarian-regime dummies and, hence, are pre-treatment: ethnic and religious fractionalization, oil resources, civil conflict, and militarized interstate disputes.

On one hand, however, one could argue that at least the conflict controls induce post-treatment bias as, for example, personalist dictators are more likely to be involved in domestic or interstate conflict. Hence, the main text also presents analyses that omit all substantive controls (and only include the fixed effects next to the temporal controls). On the other hand, it may be an effort worth making to examine the impact of post-treatment bias more thoroughly. To this end, we have decided to include two additional controls that seem likely to be influenced as a consequence of our treatments: GDP per capita and population. These controls have been identified as major determinants of domestic (in)stability, but may also be affected by, e.g., personalism. The national income and population size may decrease once personalists have assumed power, for example. Both variables are taken from the World Bank's World Development Indicators and are log-transformed.

Appendix Table 8. Post-Treatment Bias

	Appendix Model 18	Appendix Model 19
Single-Party Regime	-0.451	
	(0.709)	
Military Regime	-1.076	
	(1.071)	
Monarchy	-1.139	
	(1.223)	
Personalist Regime		0.641
		(0.743)
Ethnic Fractionalization	-0.772	-1.278

	(1.065)	(1.013)
Religious Fractionalization	0.526	0.998
	(2.134)	(2.135)
Oil	0.162	0.219
	(0.561)	(0.560)
Civil Conflict	-0.414	-0.371
	(0.508)	(0.489)
Militarized Interstate Dispute	-0.099	-0.150
	(0.265)	(0.272)
GDP per capita (ln)	-0.068	-0.049
	(0.243)	(0.248)
Population (ln)	-0.035	-0.033
	(0.136)	(0.135)
Constant	5.427	4.716
	(3.620)	(3.939)
Obs.	1, 281	1, 281
Region Fixed Effects	Yes	Yes
Wald χ^2	318.11	233.03
Prob > χ^2	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on country in parentheses; region fixed effects and temporal controls included, but not reported.

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

We have included both variables into our main models (Models 4 and 7 of the main text) as well as Appendix Models 16 and 17, which are based on the counterbalancing data from Pilster and Böhmelt (2011, 2012). While the inclusion of these items in Models 16-17 does not alter the substance of our main finding, it does so in Models 18 and 19. Specifically, when including the income and population controls, the authoritarian-regime dummies are no longer statistically significant at conventional levels, although the direction of the impact is as expected. However, the two controls as such do not add much to the explanatory power of the models either as they are insignificant. Eventually, it seems more likely that they do indeed soak up some of the impact of the treatment variables, as we have induced post-treatment bias by incorporating these additional controls.

Appendix Table 9. Poisson Models

	Appendix Model 20	Appendix Model 21
Single-Party Regime	-0.270	
	(0.107)**	
Military Regime	-0.374	
	(0.187)**	
Monarchy	0.032	
	(0.144)	
Personalist Regime		0.244
		(0.100)**
Ethnic Fractionalization	-0.193	-0.058
	(0.343)	(0.330)
Religious Fractionalization	0.785	0.477
	(0.493)	(0.420)
Oil	0.095	0.057
	(0.111)	(0.111)
Civil Conflict	0.094	0.054
	(0.087)	(0.084)
Militarized Interstate Dispute	-0.002	0.006
	(0.050)	(0.051)
Constant	0.836	0.702
	(0.253)***	(0.240)***
Obs.	1, 810	1, 810
Region Fixed Effects	Yes	Yes
Wald χ^2	128.36	131.08
Prob > χ^2	0.000	0.000

Note: Table entries are Poisson regression coefficients; standard errors clustered on country in parentheses; region fixed effects and temporal controls included, but not reported.

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

Appendix Table 10. Year Fixed Effects

	Appendix Model 22	Appendix Model 23
Single-Party Regime	-0.412	
	(0.321)	
Military Regime	-1.521	
	(0.451)***	
Monarchy	0.375	
	(0.515)	
Personalist Regime		0.602
		(0.329)*
Ethnic Fractionalization	0.129	-0.048
	(0.564)	(0.503)

Religious Fractionalization	0.675	0.205
	(1.028)	(1.104)
Oil	-0.401	-0.519
	(0.326)	(0.331)
Civil Conflict	0.054	-0.068
	(0.266)	(0.246)
Militarized Interstate Dispute	-0.135	-0.155
	(0.235)	(0.229)
Constant	-1.111	-1.114
	(0.939)	(0.775)
Obs.	1,810	1,810
Year Fixed Effects	Yes	Yes
Wald χ^2	405.68	420.43
Prob > χ^2	0.000	0.000

Note: Table entries are logistic regression coefficients; standard errors clustered on year in parentheses; region fixed effects, year fixed effects, and other temporal controls included, but not reported.

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

Count Models

Due to various reasons stated in the main text, we employ a binary indicator for our outcome of interest, i.e., counterbalancing. However, De Bruin (2018) also offers a variable on the *number* of military organizations with bases within 60 miles of the capital having access to centers of power. Border and frontier guards and rural militia are excluded for this variable, while police forces are included if they are militarized. We re-estimated our main models using this count variable as the outcome of interest to examine the robustness of our results. Due to the different dependent variable, we also replace the logistic regression models by Poisson models. Appendix Table 9 presents our results, but the results are virtually identical to those discussed in the main text as demonstrated in Appendix Models 20 and 21.

Year Fixed Effects

In a last robustness check, we examined the importance of common temporal shocks. Common global trends affect all states simultaneously in the system and examining this aspect of

temporal dependence. Appendix Table 10 thus adds year fixed effects. Examining the joint statistical significance of the battery of yearly dummies shows that at least one of their coefficients is different from 0, which suggests that common exposure does exist. However, our main findings remain unaltered.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110(3): 512-529

Barany, Zoltan D. 1991. "Civil-Military Relations in Communist Systems: Western Models Revisited." *Journal of Political and Military Sociology* 19(1): 75-99.

Bienen, Harry, and Nicolas Van de Walle. 1989. *Of Time and Power: Leadership Duration in the Modern World*. Stanford, CA: Stanford University Press.

Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(2-1): 67-101.

De Bruin, Erica. 2018. Preventing Coups d'etat: How Counterbalancing Works. *Journal of Conflict Resolution* 62(7): 1433-1458.

Elwert, Felix, and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology* 40(1): 31-53.

Herspring, Dale R. 1999. "Samuel Huntington and Communist Civil-Military Relations." *Armed Forces & Society* 25(4): 557-577.

Geddes, Barbara, Joseph Wright, and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions: A New Data Set." *Perspectives on Politics* 14(1): 313-31.

King, Gary, and Langche Zeng. 2006. "The Dangers of Extreme Counterfactuals." *Political Analysis* 14 (2): 131–159.

Marshall, Monty and Keith Jaggers. 2013. *POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2002. Dataset User's Manuel*. College Park, MD: University of Maryland.

Narang, Vipin and Caitlin Talmadge. 2017. Civil-Military Pathologies and Defeat in War: Tests Using New Data. *Journal of Conflict Resolution*: forthcoming.

Perlmutter, Amos, and William M. LeoGrande. 1982. "The Party in Uniform: Toward a Theory of Civil-Military Relations in Communist Political Systems." *American Political Science Review* 76(4): 778-789.

Pilster, Ulrich and Tobias Böhmelt. 2011. "Coup-Proofing and Military Effectiveness in Interstate Wars, 1967-1999." *Conflict Management and Peace Science* 28(4): 331-350.

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): 355-72.

Plümper, Thomas and Vera E. Troeger. 2011. "Fixed-Effects Vector Decomposition: Properties, Reliability, and Instruments." *Political Analysis* 19(2): 147-164.

Powell, Jonathan and Clayton Thyne. 2011. "Global Instances of Coups from 1950-2010: A New Dataset." *Journal of Peace Research* 48(2): 249-259.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.

Talmadge, Caitlin. 2016. "Different Threats, Different Militaries: Explaining Organizational Practices in Authoritarian Armies." *Security Studies* 25(1): 111-141.

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.