# Online Appendix: Personalization of Power and Repression in Dictatorships

Erica Frantz*[1], Andrea Kendall-Taylor[2], Joseph Wright†[3], and & Xu Xu‡[3]

[1]Michigan State University
[2]Center for New American Security
[3]Pennsylvania State University

November 16, 2018

**Abstract**

This article uses new data measuring gradations of personalism in authoritarian regimes to evaluate the relationship between concentration of power and patterns of repression. It shows that the personalization of power in dictatorships leads to an increase in repression. Given the rise in personalism we are witnessing globally, the findings of this study imply that repression is likely to become more prevalent in dictatorships as a consequence.

Accepted for publication at *Journal of Politics*

*ericaemilyfrantz@yahoo.com
†josephgwright@gmail.com
‡xux112@psu.edu

## Non-essential or supplementary footnotes

Given space restraints in a short article, we place some of the citations and footnotes here, rather than in the main text. Footnotes that remain in the main text respond directly to reviewer comments.

### Introduction

- The following sentence could have an additional citations:

  "High-level Chinese officials are still vetted by the party apparatus, and – while loyalty to Xi is important for promotion – competence still remains a key criterion (Jia and Xu, 2018: 3; Shirk, 2018: 25)."

### Empirical approach

- The data in the personalism index are coded for January 1 of each calendar year, which means the change occured in the *prior* calendar year, effectively lagging the variable.
- The final item, purge of military/security officers, may be conceptually close to the dependent variable. The Appendix shows that the results remain when omitting the purge variable from the measure of personalism.
- The share of with-case variation is 36 percent, a number similar across all regime *types*.

### Hypothesis-testing models

- Estimating a regime-case fixed effects linear model means the empirical approach accounts for all differences between authoritarian regimes, including how the regime seized power, (time invariant) state strength, colonial history, autocratic regime type, and geographic region.
- We report clustered standard errors; the Appendix lists variable sources and definitions.

### Predictive models of repression

- The Appendix shows that a random effects (RE) estimator with a binary indicator for the Cold War has substantially smaller prediction error than the two-way fixed FE models analyzed above. This suggests two-way FE models over-fit the data. The Appendix reports RE models similar to those in Figure 1, with stronger personalism estimates.
- We report the root mean squared error (RMSE) to assess predictive accuracy. The baseline RMSE varies for each test variable because samples differ due to data availability.
- Comparing results in Figure 1(b) with those in Hill and Jones (2014), we note two differences. Our sample extends from 1950 to 2010, theirs 1980 to 2000. Second, our sample includes only dictatorships, not dictatorships *and* democracies. Variables that differ between democracies and dictatorships, such as Polity, may not predict repression in our sample as well as in theirs.

- Footnote 2 with citations is: The low explanatory power of institutional variables may reflect the possibility that larger structural features of regime's historical political economy explain both the choice of institutions and repression (Pepinsky, 2014) or the the possibility that institutions operate differently in different kinds of regimes (see e.g. Wilson and Wright, 2017).

## Appendix S: Data description

Table S-1 shows the summary statistics for the data and sample employed in the models in Figure 1 in the main text. All continuous variables have been standardized to have a mean of 0 and a standard deviation of 1. The sample period for models in Figure 1 extends from 1955 to 2010 because this is the coverage period for the protest variable. The missing data for GDP per capita and population are for Singapore.

Table S-1: Summary statistics for Figure 1

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Repression | 0.016 | 1.003 | -3.416 | 2.817 | 4214 |
| Year | 1982.168 | 14.766 | 1955 | 2010 | 4214 |
| Personalism | 0.426 | 0.277 | 0 | 1 | 4214 |
| Leader time in power (logged) | 0.026 | 0.991 | -1.831 | 2.123 | 4214 |
| GDP per capita (logged) | 0 | 1 | -3.364 | 3.754 | 4164 |
| Population (logged) | 0 | 1 | -2.577 | 3.635 | 4164 |
| Civil conflict | 0.033 | 0.178 | 0 | 1 | 4214 |
| Int'l conflict | 0.027 | 0.162 | 0 | 1 | 4214 |
| Protest | 0 | 1 | -2.746 | 2.756 | 4194 |
| Senior officer | 0.191 | 0.393 | 0 | 1 | 4214 |
| Junior officer | 0.16 | 0.367 | 0 | 1 | 4214 |
| Institutions | 0.438 | 0.383 | 0 | 1 | 4214 |
| Year | 1982.168 | 14.766 | 1955 | 2010 | 4214 |

Figure S-1 shows the distributions of the two main variables. The left panel shows the repression level as a standardized variable with mean of 0 and a standard deviation of 1. We rescaled the latent personalism index so that its minimum observed value is 0 while its maximum observed value is 1. We do this to ease interpretation of this variable in a linear model, particularly to facilitate visual comparison of estimates of personalism and binary variables, such as civil conflict, international conflict, senior officer, and junior officer.

**Data description** Tables S-2 and S-3 list the coverage years and sources for the data used in Figures 1 and 2, respectively. Violent and non-violent protest are binary indicators of an ongoing anti-regime protest campaign. Ethnic exclusion is the lagged share of the ethnically relevant population excluded from executive power. Coup attempt is a binary indicator of a coup attempt in the current or prior two years. Election period is an indicator of a multiparty election (opposition allowed, more than one party legal, choice of candidates on the ballot) in the current, prior, or subsequent year. Youth bulge is the ratio of the youth population (15-24 years) relative to the total adult population (15 and older). Two variables may not be familar to readers: divided military seizure group and inherited party. The former is a proxy for whether the military that seized power is divided (either a junior officer coup – implying some senior officers did not join coup plot – or an unarmed uprising – implying the military did not successfully violently repress the uprising). The latter measures whether a supporting political party existed prior to the regime seizing power. We include these as test variables because Geddes, Wright and Frantz (2018) show that these pre-seizure features of autocracies are highly correlated with the personalist index. Thus, the

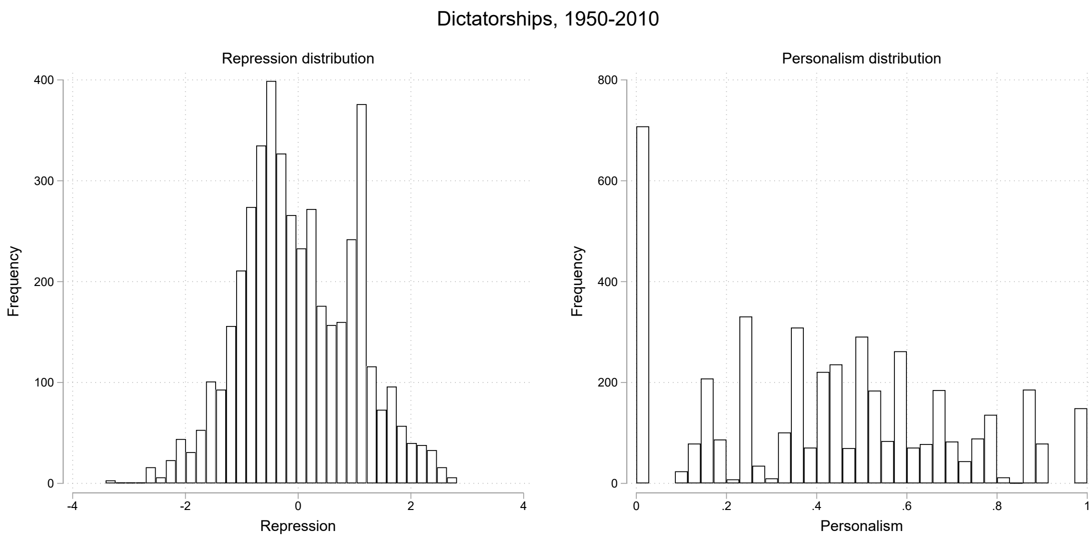


Figure S-1: Distributions of key variables.

predictive power of the personalism might simply be an artifact of these pre-seizures features of autocracies and not changes in leader behavior that constitute increasing personalism. We show in the cross-validation analysis that this is not the case.

Table S-2: Figure 1 data

| Variable | Description | Years | Source |
|---|---|---|---|
| Repression | latent; continuous; standarized | 1950–2010 | Fariss (2014) Vers. 2 |
| Personalism | latent; 0-1 scale | 1950–2010 | GWF (2018) |
| GDP per capita | log transform; standarized | 1950–2010 | EPR Vers. 3 (PWT, WDI) |
| Population | log transform; standarize | 1950–2010 | EPR Vers. 3 (PWT, WDI) |
| Leader time | log transform; standarized | 1950–2010 | GWF (2018) |
| Civil conflict | binary; high intensity | 1950–2010 | Gleditsch et al. (2002) |
| Intern'l conflict | binary; high intensity | 1950–2010 | Gleditsch et al. (2002) |
| Protest | latent; continuous; standarized | 1955–2010 | See protest data section below |
| Senior officer | binary | 1950–2010 | GWF (2018) |
| Junior officer | binary | 1950–2010 | GWF (2018) |
| Institutions | eight ordinal levels; 0-1 scale | 1950–2010 | GWF (2018) |

EPR: Wimmer, Cederman and Min (2009); GWF (2018): Geddes, Wright and Frantz (2018)

Table S-3: Figure 2 data for test variables

| Variable | Years | Source |
| --- | --- | --- |
| Violent protest | 1950–2006 | NAVCO 2 |
| Non-violent protest | 1950–2006 | NAVCO 2 |
| Anocracy | 1950–2010 | EPR Vers. 3 (Polity2) |
| Leader age | 1950–2010 | GWF (2018) |
| Ethnic exclusion | 1950–2010 | EPR Vers. 3 |
| Coup attempt | 1950–2010 | Powell and Thyne (2011) |
| Party regime | 1950–2010 | GWF (2014) |
| Military regime | 1950–2010 | GWF (2014) |
| Monarchy | 1950–2010 | GWF (2014) |
| Personalist regime | 1950–2010 | GWF (2014) |
| Economic growth | 1950–2010 | EPR Vers. 3 (PWT, WDI) |
| Election period | 1950–2010 | NELDA 4 |
| Support party | 1950–2010 | GWF (2018) |
| Prior democracy | 1950–2010 | GWF (2018) |
| Oil rents (log) | 1950–2010 | EPR Vers. 3 |
| Divided seizure group | 1950–2010 | GWF (2018) |
| Inherited party | 1950–2010 | GWF (2018) |
| Polity score | 1950–2010 | EPR Vers. 3 (Polity2) |
| Youth bulge | 1950–2000 | Urdal (2006) |
| Judicial independence | 1950–2010 | Linzer and Staton (2015) |

NAVCO 2: Chenoweth and Lewis (2013); EPR: Wimmer, Cederman and Min (2009); GWF (2014): Geddes, Wright and Frantz (2014, 2018); NELDA 4: Hyde and Marinov (2012)

**Protest data** The protest data come from Chenoweth, D'Orazio and Wright (2014). In this project, the authors use information from eight existing data sets that measure anti-government protest cross-nationally. Table S-4 lists the eight datasets, the geographic and temporal coverage of each, as well as the type of data collected in each. The raw data sets include country-year counts of protest levels (Banks), event data with daily information from news reports (e.g. ACLED, SCAD, and SPEED), and campaign data that measures long-term protest campaigns that can last for a couple of weeks up to multiple years (MEC). The latter, for example, includes the three-week Georgian Rose Revolution protests in November 2003 as well as the six-year anti-Pinochet campaign in Chile that started with the May 1983 National Protest[1] and ended with the 1989 transition to civilian rule.

Table S-4: Data sets used to construct latent protest variable

| Data set | Temporal coverage | Spatial coverage | Data type |
|---|---|---|---|
| ACLED | 1997-2013 | Africa | daily event |
| SCAD | 1990-2011 | Africa | event |
| ECPD | 1980-1995 | Europe | daily event |
| SPEED | 1950-2012 | Global | daily event |
| LAPP | selected years | Latin Am. | daily event |
| IDEA | 1990-2004 | global | daily event |
| Banks | 1950-2012 | global | country-year count |
| MEC | 1955-2013 | global | campaign |

Armed Conflict Location and Event Dataset (acled)
Downloaded from: `https://www.strausscenter.org/acled.html` on 9.12.13.
Version: ACLED All Africa 1997-November 2013.
Data structure: daily event; each row records event that occurs for no longer than 1 day

Cross-National Time-Series Data Archive (banks)
Downloaded from: `www.databanksinternational.com` on 9.1.13.
Version: data available on retrieval date
Data structure: country-year

European Protest and Coercion Data from Ronald Francisco (epcd)
Downloaded from: `http://web.ku.edu/ ronfrand/data/index.html` on 9.1.14.
Version: data available on download date.
Data structure: daily event; each row records event that occurs for no longer than 1 day

[1]Garretón (1988: 11-12) writes that "[t]he first massive demonstration, known as the National Protest, occurred in May of 1983. The Copper Workers' Confederation (CTC) had initially called for a National Strike. However, a few days beforehand they decided instead to call for a broad-based protest."

Integrated Data for Event Analysis (idea)

Downloaded from: http://thedata.harvard.edu/dvn/ on 2.12.13.
Version: http://hdl.handle.net/1902.1/FYXLAWZRIA UNF:3:dSE0bsQK2o6xXlxeaDE
IQSS Dataverse Network [Distributor] V3 [Version]
Data structure: daily event; each row records event that occurs for no longer than 1 day

Latin American Political Protest Project (lapp)

Downloaded from: http://faculty.mwsu.edu/politicalscience/steve.garrison/LA
on 9.12.13.
Version: data available on download date.
Data structure: daily event; each row records event that occurs for no longer than 1 day

Major Episodes of Contention Data Project (mec)

Obtained from Erica Chenoweth on 4.1.14.
Version: MEC Cat4 1950-2013.
Data structure: event; each row records event that occurs for multiple days to years

Social Conflict in Africa Database (scad)

Downloaded from: https://www.strausscenter.org/scad.html on 9.12.13.
Version: SCAD 3.0 1990-2011.
Data structure: event; each row records event that occurs for possibly multiple days

Social, Political and Economic Event Database Project (speed)

Downloaded from: http://www.clinecenter.illinois.edu/research/speed-data.ht
on 2.12.13.
Version: data available on retrieval date
Data structure: daily event; each row records event that occurs for no longer than 1 day

**Dynamic IRT model** The item response theory (IRT) model[2] combines information from multiple data sets to estimate a latent mean value of protest at the country-year level. The IRT model used in an updated approach is dynamic in the treatment of the item-difficulty cut-points of the latent variable and employs a negative binomial distribution to model count data (rather than binary data) in the items. The resulting data set has global coverage for the period from 1955 to 2010.

[2]The item response theory (IRT) approach used in the paper allows the authors to combine information from multiple sources that may not overlap in their temporal and spatial coverage. This approach thus circumvents missing data issues that arise from other measurement approaches, such as clustering and factor analysis, that use listwise (row) deletion to obtain a rectangular data object for estimating a latent variable.

## Appendix A: Additional results

Figure A-1 reports models that add covariates to the final specification in Figure A-1, one at a time. We add the following covariates: leader age; violent anti-regime protest campaign (sample period ends in 2007); non-violent protest campaign (sample period ends in 2007); judicial independence; multiparty election; recent past coup; and youth bulge. In all these specifications, the estimate for personalism remains stable, positive, and statistically significant.

The models reported in Figures 1 and A-1 are two-way fixed effects models where the cross-section unit is the autocratic regime-case. (There are 462 leaders nested in 261 regime-cases nested in 119 countries from 1955-2010.) We cluster standard errors on leaders, reflecting the fact that the personalism measure is coded using observed *leader* behavior, such as whether the leader creates a personally loyal security organization or establishes a new support party. Modeling a fixed regime cross-section unit accounts for all time-invariant differences between autocratic regimes: geography; colonial history; prior regime type; time-invariant ethnic (ex)inclusion; regime type (including whether Geddes, Wright and Frantz (2014) code the regime as personalist, military, or party); state capacity; deep historical political economy differences; and how the regime seized power. Because we use a regime-case FE estimator we cannot include binary indicators of autocratic regime type (e.g. military regime, party regime, and personalist regime) in the specification.

Figure A-2 reports results from models that change the cross-section unit as well as whether the cross-section unit is modeled as a random- or fixed-effect. We examine results from three types of cross-section units: country (116-119 units), regime-case (258-261 units), and leader (456-462 units).[3] To illustrate the differences in cross-section units, we discuss autocracies in Iran. The country-effects models treat Iran, the country, as the cross-section unit. The regime-case effects models distinguish between the monarchical Pahlavi regime and the theocratic regime that came to power during the 1979 Revolution. The first regime-case, the Pahlavi regime, had only one leader during the sample period, while the second regime-case has had two leaders, Khomeini and Khamenei. Therefore, there is one country unit for Iran, two regime-case units, and three leader units.

The results for personalism reported in Figure A-2 indicate that changing the cross-section unit or changing whether we model the cross-section unit as a random or fixed effect does not appreciably change the estimate of interest. This should not be entirely surprising because the latent measure of personalism is coded to capture changes over time in observed leader behavior that Geddes, Wright and Frantz (2018) argue are plausible manifest indicators of increasing consolidation of power in the leader's hands. We refer readers to Geddes, Wright and Frantz (2017) for more information on the time-varying measure of personalism, but highlight that this variable contains substantial *within-leader* and *within-regime* variation.

Figure A-3 reports results from random effects (RE) models that mirror the fixed effects (FE) models reported in Figure 1. The estimates for personalism in the RE models are larger than those in the FE models reported in Figure 1 in the main text. If we believe the FE estimator overfits the data – as suggested by the cross-validation analysis – the results in A-3 suggest that the FE estimator may also bias the estimates of personalism towards `0`.

Focusing on the final models reported in Figures 1 and A-3, we find that a Hausman test rejects the null that RE and FE estimators yield similar estimates for the explantory variables. However,

[3] In all these models, we estimate standard errors clustered on the same cross-section unit as the unit effect.


Figure A-1: Additional covariates. Two-way fixed effects, with clustered standard errors.

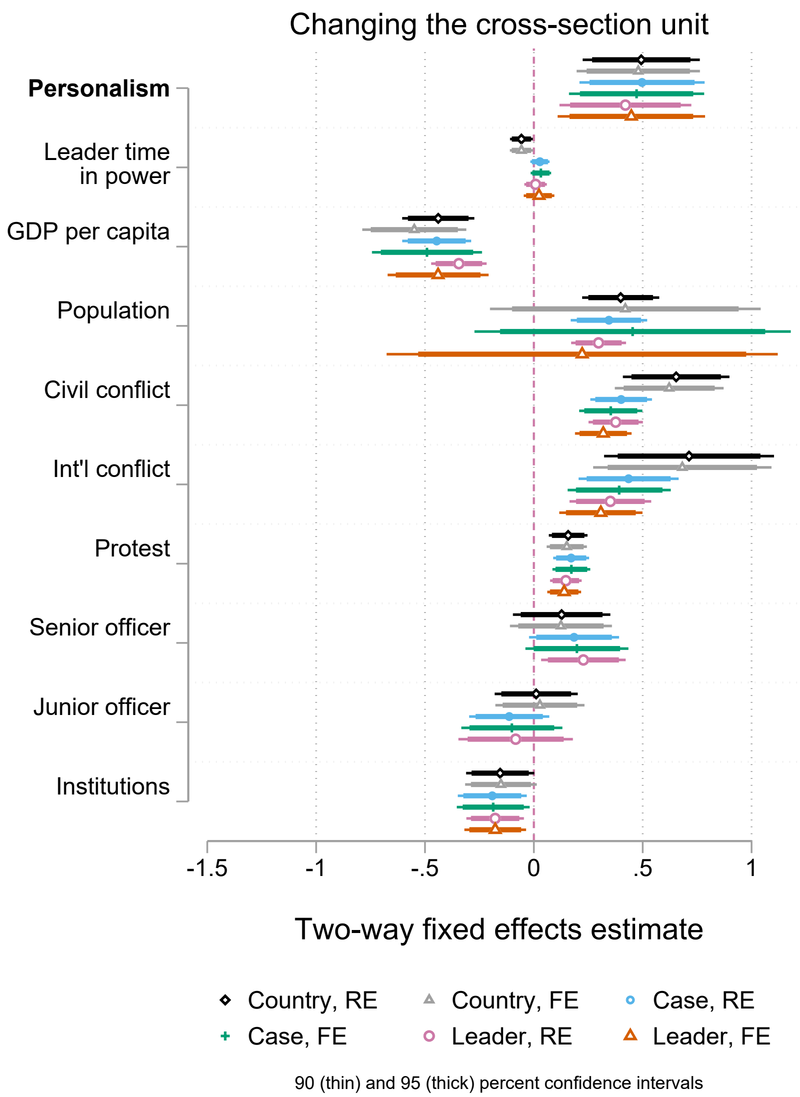


Figure A-2: Changing the cross-section unit. Random and fixed effects models, with clustered standard errors.

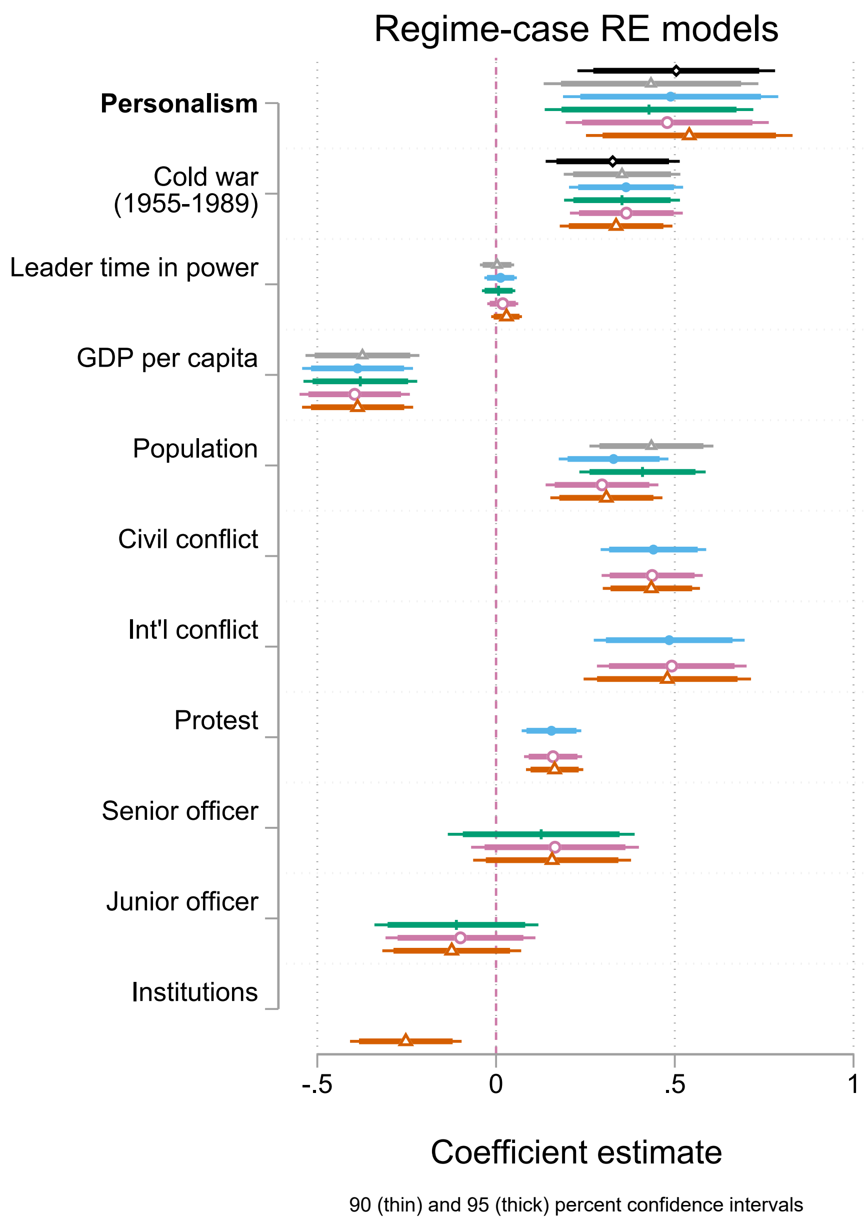


Figure A-3: Reported models with RE with clustered standard errors.

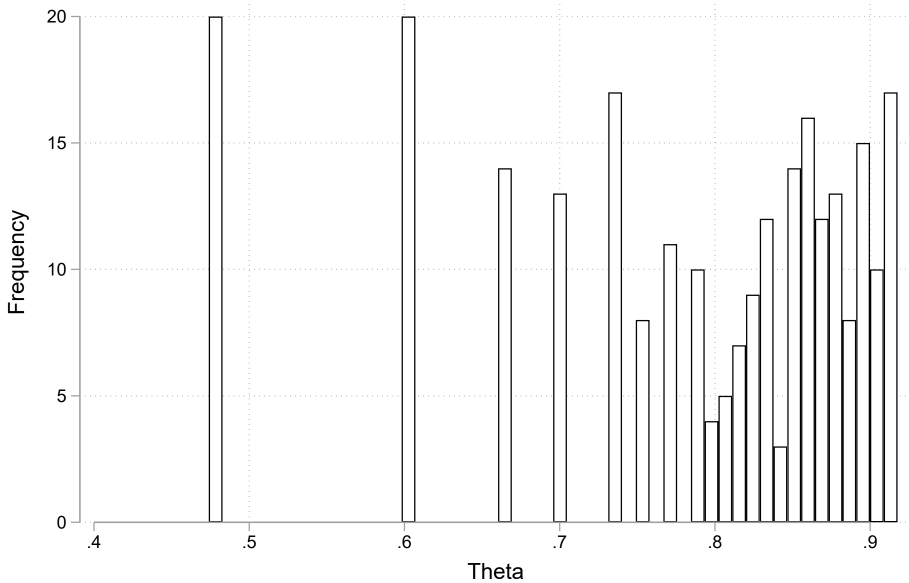


Figure A-4: Distribution of $\theta$'s in the RE model. One $\theta$ estimate reported for each of the 258 panels.

the differences in the estimates for personalism across the two estimators are small and *not* statistically significant (0.529 vs. 0.541). Figure A-4 plots the distribution of $\theta$'s from the RE model. We note that most of the $\theta$'s are greater than 0.75, meaning that for these panels the 'weights' assigned to the group means are close to 1, which is equivalent to the fixed effects estimator. This visualizes the intuition for why – in this application – the RE and FE estimators yield similar results for variables, such as personalism, that have non-trivial within-unit variation.

Figure A-5 shows the results from a series of error-correction models that estimate the dynamic relationship between personalism and repression. The model is the following, where $P$ is personalism, $X$ is a vector of covariates, $\nu_t$ are year effects, and $\epsilon_{it}$ is the error term:

$$\Delta Repression = Repression_{t-1} + \Delta P + P_{t-1} + \Delta X + X_{t-1} + \nu_t + \epsilon_{it}$$

While this equation separately estimates the short- and long-term effects of yearly changes in personalism, we report the long-run multiplier (total long run effect) using the Bewely transformation. In all but one specification the long-run multiplier for personalism is positive and significant, indicating that once we model temporal dynamics, personalism is statistically associated with increased repression. The one specification where the estimate of the long-run multiplier for personalism is only significant at the 0.10 level (i.e. not at the 0.05 level) is the specification that omits conflict variables (civil war, international war, and domestic anti-regime protest).

In the cross-validation exercise reported the main text (Figure 2) we note that the two-way FE model overfits the data, and that an RE (in lieu of FE) model with a Cold War indicator variable (instead of year effects) fits the data better. Figure A-6 reports the RMSE sampling distribution from 10-fold cross-validation for three estimators, each with the baseline model specification that only includes two covariates: GDP per capita and population. The vertical axis depicts the RMSE. The point estimates are the median change in prediction error and the confidence bands are the 2.5

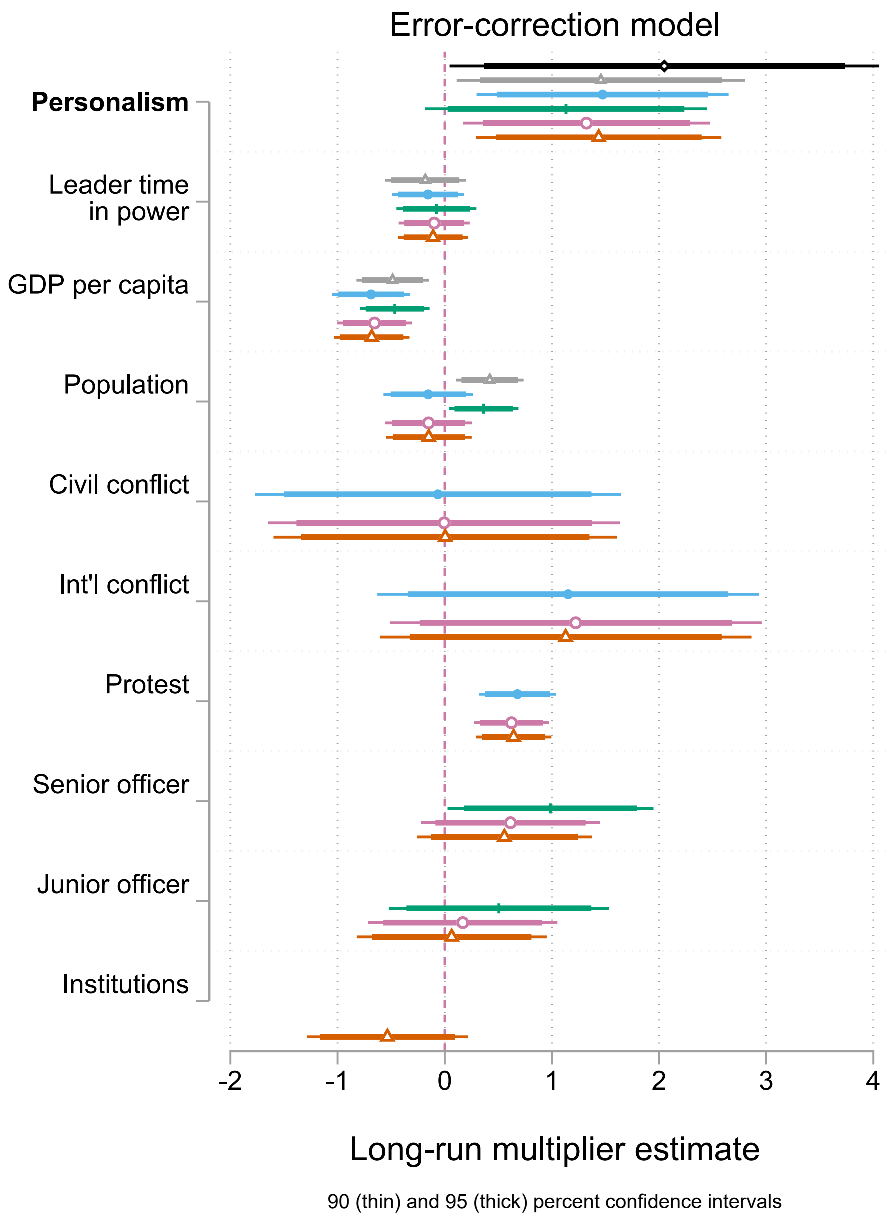


Figure A-5: Error-correction models with clustered standard errors.

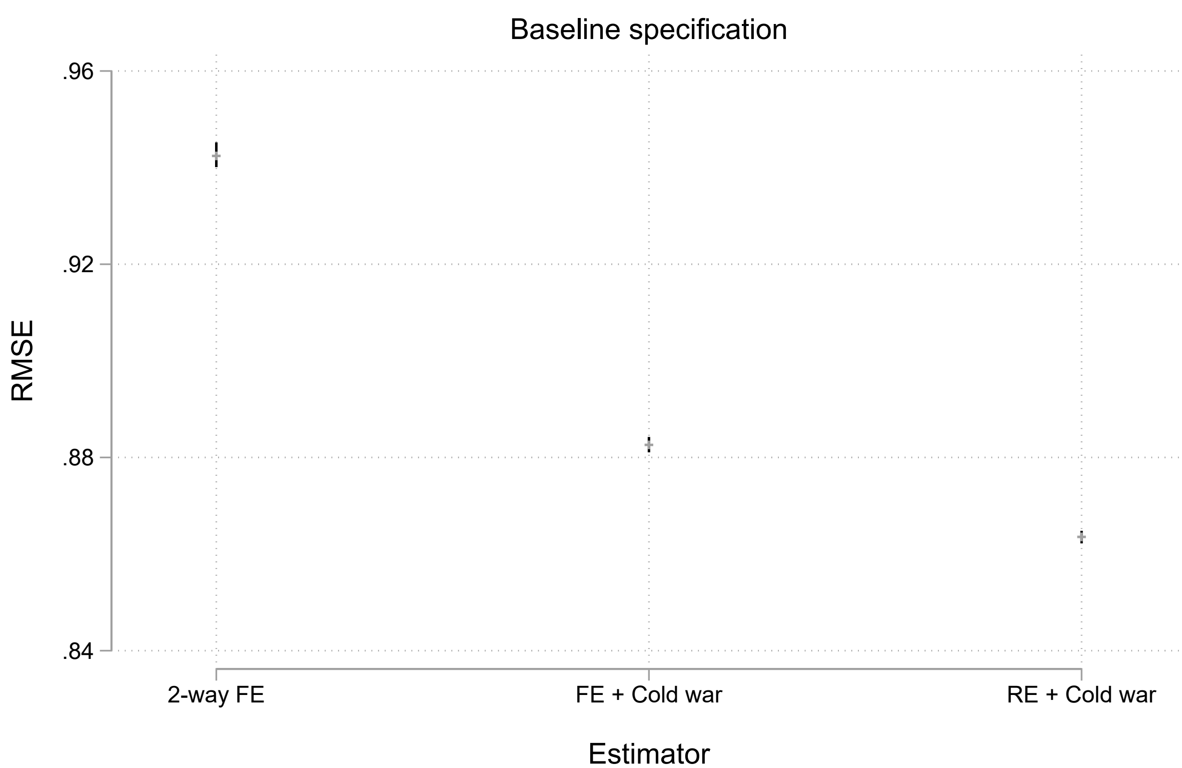


Figure A-6: RMSE for different estimators.

and 97.5 percentiles of the sampling distribution of the RMSE statistic from 1000 simulations. The RMSE for the two-way FE model (i.e. that reported in Figure 1) is over 0.94. Removing the year effects and modeling the common time trend with a Cold War dummy lowers the RMSE to just over 0.88. Using an RE estimator instead of the FE estimator lowers the RMSE still further. For this reason, we use an RE estimator with a Cold War indicator instead of the two-way FE estimator in the cross-validation analysis.

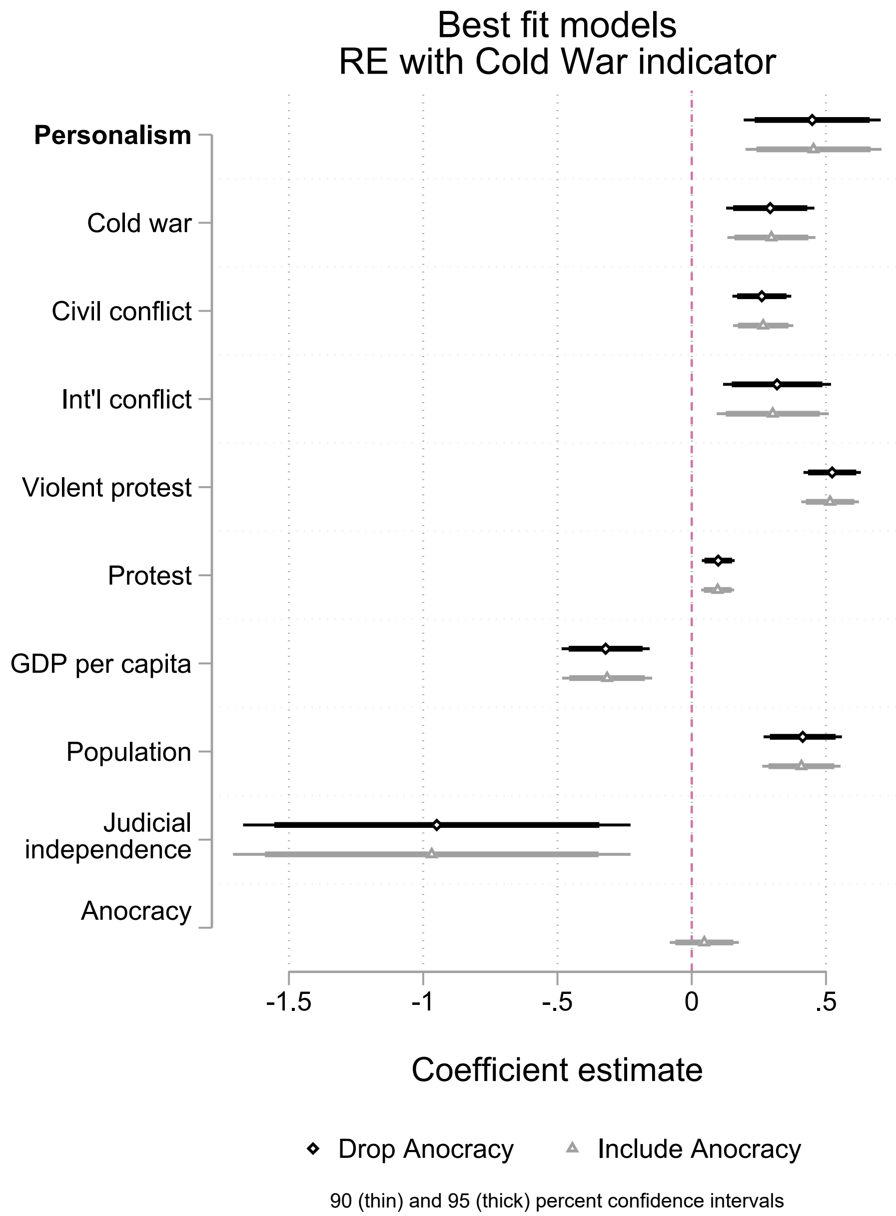


Figure A-7: Best fit models with clustered standard errors.

Figure A-7 reports the coefficient estimates from a model that best fits the data, given the results of the cross-validation analysis reported in Figure 2 in the main text. The first model specification includes all the covariates that lower the RMSE more than personalism (conflict variables, protest, GDP per capita, population, and judicial independence) in addition to personalism. The second specification adds anocracy. This 'best fit' model yields a similar but slightly smaller estimate for personalism (0.449) than the estimate in the full specification in Figure 1 in the main text (0.472).

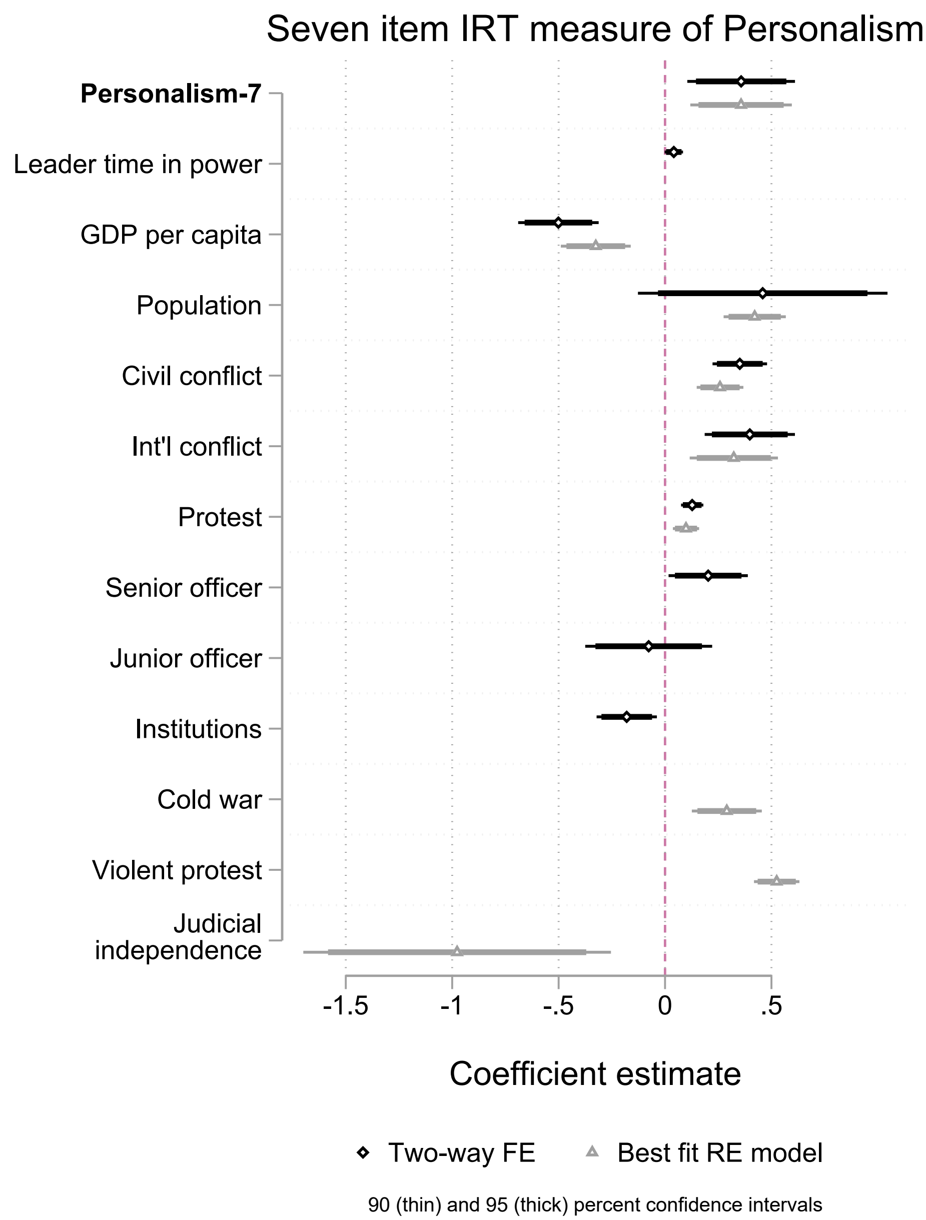


Figure A-8: Seven-item IRT model to measure personalism.

The analyses in the main text utilize the personalism index from Geddes, Wright and Frantz (2018), described in detail in Geddes, Wright and Frantz (2017). This index contains eight manifest indicators of personalism. One, however, is perhaps too conceptually and operationally close to repression to be used as an explanatory variable for repression: purge of senior military or security officers. A military purge may constitute an observed instance of repression if the regime jails, tortures, or kills the purged officers. To ensure that this item – and its informational contribution to the personalism index – is not driving the main result, we re-calculated the personalism index using only the seven other items. The two personalism indices are correlated at 0.97. We then re-tested the final specification of the two-way FE model (see Figure 1) and the best fit RE model (see Figure A-7). Figure A-8 reports these tests. In both models, the main result for personalism remains, indicating that the reported results do not rely on information from possibly repressive military purges.

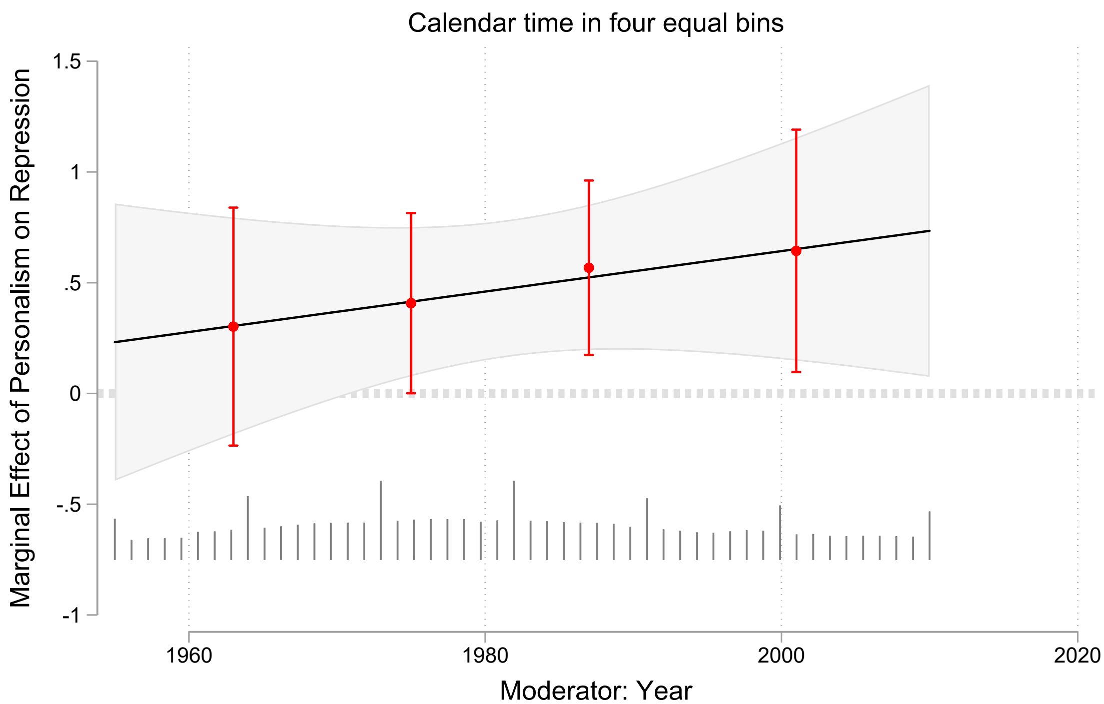


Figure A-9: Marginal effect of personalism across calendar time.

Figure A-9 plots the average pointwise marginal effect of personalism from the two-way fixed effects model with covariates (the last model specification reported in Figure 1 in the main text). The plot reports the average estimated pointwise marginal effects for calendar years within four separate time periods, each with an equal number of observations. This allows us to investigate whether the marginal effect of personalism varies over time period. In general, the statistical association between personalism and repression is slightly increasing over time. The average pointwise marginal effects within each of the four bins is close to the line (in black) plotting the varying marginal effects of personalism obtained from a linear interaction between calendar year and personalism. We include this plot to verify that the (positive) marginal effect of personalism persists in the post-Cold War period.

## Appendix B: Modeling uncertainty in the latent dependent variable

Figures B-1 and B-2 show results from models that incorporate the uncertainty in the dependent variable, latent repression, as well as the uncertainty in the main explanatory variable, latent personalism. Figure B-1 shows results similar to the final specification reported in Figure 1, using a two-way fixed effects estimator with covariates and clustered standard errors. Figure B-2 shows the result from the best-fit model specification, which is the random effects estimator with a Cold War dummy variable. We generate 1000 draws for each country-year observation of repression and separately for personalism from a normal distribution described by the mean and standard deviation of the latent variable estimate for each country-year observation of these variables. We then estimate the model 1000 times, for each draw from the distributions of the latent variables (Crabtree and Fariss, 2015). The reported point estimates are the mean of the point estimates from these 1000 models, while the reported variance is the sum of the between variance, the within variance, and the sampling variance, per Rubin (1987).

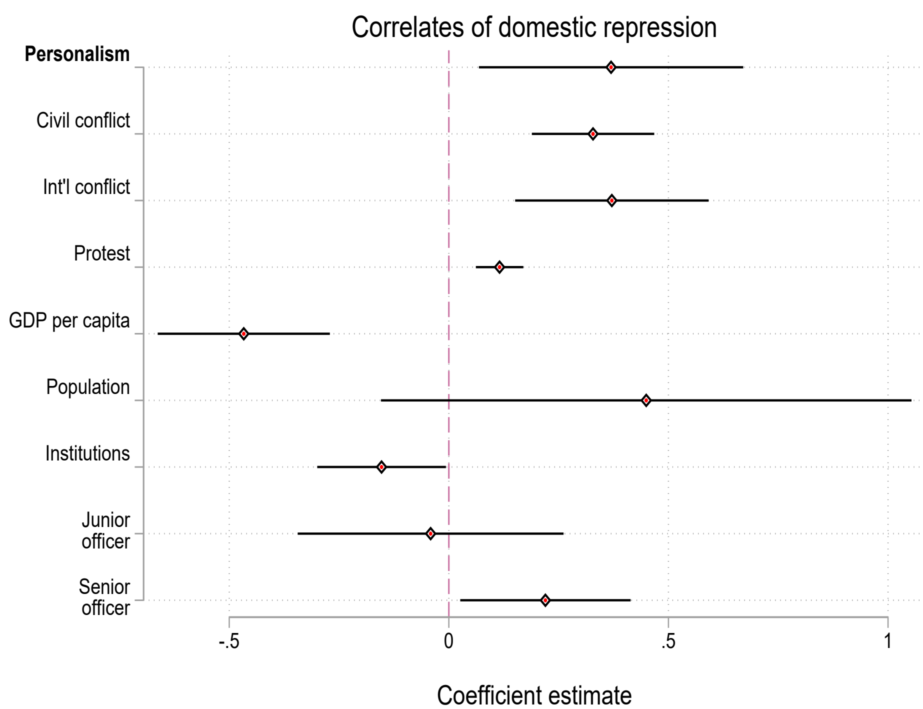


Figure B-1: Modeling uncertainty in repression and personalism variables. Two-way fixed effects, with clustered standard errors.

In both model specifications, the point estimate for personalism is similar to that reported in the main text. However, the estimated variance of this estimate is slightly larger. That said, in both models the estimate of theoretical interest is statistically different from zero at the 0.05 level.

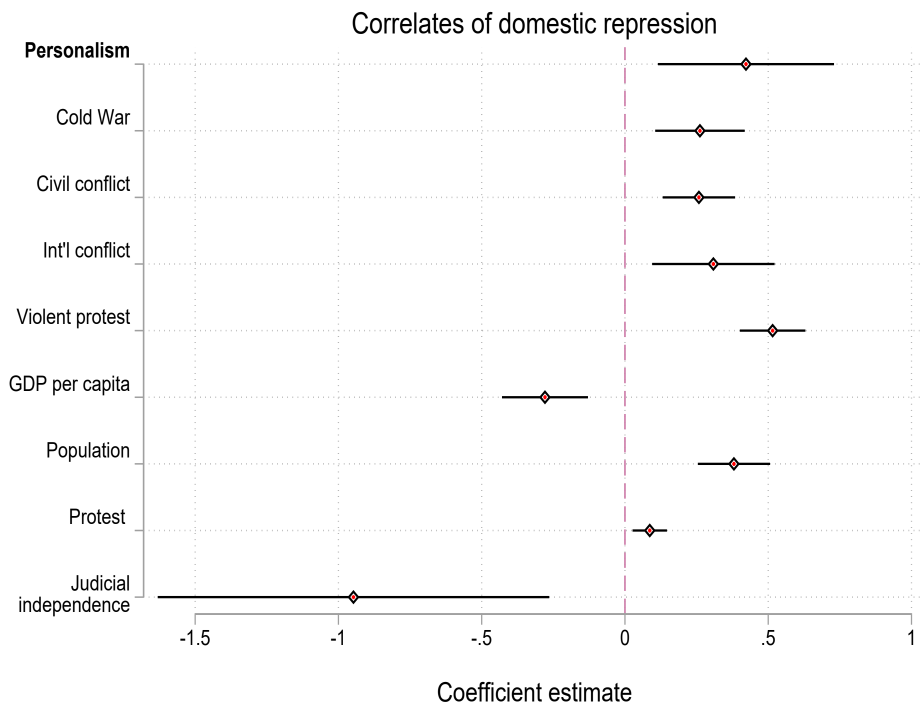


Figure B-2: Modeling uncertainty in repression and personalism variables. Random effects best fit specification, with clustered standard errors.

## Appendix C: Bounding bias from selection on unobservables

This Appendix reports results from tests of the potential bias from unmodeled variables. The measure of personalism is based on observed real-world political phenomena, not random assignment of a treatment variable. Indeed, it is implausible to conceive of a treatment variable (a) that approximates real-world consolidation of state power in the hand of dictatorial leaders, and (b) that researchers could ethically and practically assign randomly. Therefore, we cannot rule out the possibility of bias in the estimate of personalism based on unmodeled variables.

Thus, following the work of Altonji, Elder and Taber (2005) and Oster (2017), we calculate the bounds on potential bias in the estimate of the main variable of interest, personalism, that could result from selection on unobservables. The test proposed by Oster (2017) uses information from changes in the point estimates and $R^2$ values derived from a model specification without any 'controls' ("uncontrolled") and from a specification that includes covariates as 'controls' ("controlled"), provided the 'control' variables are added to the specification to mitigate bias. Importantly, the test *only* provides credible bounds for causal interpretation of the estimate of interest *if* "the residual omitted variable bias after inclusion of controls is proportional to the coefficient movements and the ratio of the movement in R-squared with inclusion of the observable control to the expected movement in R-squared with the inclusion of the unobservable controls" (Oster, 2017: 2).

While we cannot directly estimate how including unobservables in the specification (in addition to the observed control variables) would influence the estimate of $\beta_{personalism}$ and the resulting $R^2$, we can make plausible assumptions about (a) the extent to which the $R^2$ would change when adding unobservables to the specification, and (b) the extent to which the resulting change in the estimate of $\beta_{personalism}$ is proportional to the change in $\beta_{personalism}$ when adding observed covariates (i.e. 'controls') to a baseline specification that only includes the treatment variable of interest.

Oster defines the parameter $\delta$ as the extent to which the unobserved variables are as important as the observed variables in producing a treatment effect that is zero. When $\delta$=1, the unobserved variables are equally important as the observed ones; $\delta>1$ implies that the unobserved variables are more important than the observed ones. Following Altonji, Elder and Taber (2005), Oster (2017: 20) suggests that assuming $\delta$=1 is an appropriate upper bound on $\delta$ because the control variables in the specification should already be selected by researchers because they believe *ex ante* these are the most important.

Oster also defines the hypothetical $R^2$ of a model with both observed and unobserved covariates as $R_{max}$. If $R_{max}$=1, this implies that all observed and unobserved covariates explain all the variation in the outcome, with no idiosyncratic (i.e. white noise) component to the outcome data generating process. The $R_{max}$ value is undefined because a model with unobserved covariates cannot be estimated. However, we can make an assumption about $R_{max}$ relative to the $R^2$ value from the regression with observed control variables. If we let $\tilde{R}$ denote the $R^2$ from a specification with the treatment and observed controls, then we can define $R_{max}= \pi\tilde{R}$, where $\pi$ is a parameter that varies in the extent to which unobserved covariates increase the $R^2$ – relative to $\tilde{R}$ – when added to the specification. Thus $\pi$=1 means that *all* unobserved covariates explain no further variation in the outcome, while $\pi$=2 means that unobserved covariates explain just as much variation in the outcome as treatment and the observed variables.

Table C-1 reports the results from three types of models: the two-way fixed effects model with covariates (i.e. controls) from Figure 1; a best-fit FE model with covariates from Figure A-6, including a Cold War dummy instead of year effects; and finally a best-fit OLS model, again with

Table C-1: Sensitivity to selection on unobservables

| | $\beta_{personalism}$ | $R^2$ | *R-max* equal to $R^2 \times 1.3$ | Bias-adjusted $\beta_{personalism} \mid \delta$=1 | $\delta \mid \beta_{personalism}$=0 |
|---|---|---|---|---|---|
| Two-way FE | | | | | |
| Uncontrolled | 0.475 | 0.021 | 0.303 | 0.402 | 5.57 |
| Controlled | 0.427 | 0.233 | | | |
| Best-fit FE | | | | | |
| Uncontrolled | 0.489 | 0.022 | 0.311 | 0.428 | 15.82 |
| Controlled | 0.444 | 0.239 | | | |
| Best-fit OLS | | | | | |
| Uncontrolled | 0.293 | 0.007 | 0.625 | 0.358 | 63.12 |
| Controlled | 0.339 | 0.481 | | | |

a Cold War indicator.[4] The first two columns report the estimate for personalism, $\beta_{personalism}$, and the $R^2$ for the uncontrolled (i.e. no covariates except personalism) and the controlled (i.e. with covariates) specification for each of the three models. In all three models, adding covariates increases the $R^2$; in the first two adding covariates decreases $\beta_{personalism}$. The third column reports the $R_{max}$ value for each specification for each model when we set $\pi$=1.3, which is the value that Oster (2017) suggests as a conservative assumption about the relative explanatory power of the unobserved variables relative to the observed ones.

The final two columns of Table C-1 report the relevant tests to evaluate the sensitivity of $\beta_{personalism}$ to selection on unobservables. The fourth column reports the bias-adjusted $\beta_{personalism}$ assuming $\pi$=1.3 and $\delta$=1 (Oster, 2017: 6). With these assumptions, the bias-adjusted estimate is smaller than the estimates from the uncontrolled and the controlled estimates but still greater than 0.4. This suggests that, under plausible and conservative assumptions about selection on unobservables, $\beta_{personalism}$ remains large and substantively meaningful. The fifth column reports the $\delta$ value under the assumption that $R_{max}$=1.3$\times\tilde{R}$ and the estimate of $\beta_{personalism}$ is zero. Recall that $\delta$ >1 implies that the unobserved variables are more important than the observed ones (including the treatment) in explaining the outcome. For the two-way FE model, the $\delta$ would have to be great than 5 for $\beta_{personalism}$ to fall to zero, implying that unobserved variables explain more than five times as much of the variation as the observed variables. Thus assuming the literature on repression has identified variables that explain at least one-sixth of observed repression, the estimate of personalism is likely positive. The $\delta$ value for the other models is substantially higher.

Finally, Figure C-1 shows how $\beta_{personalism}$ changes as we vary $\pi$, while assuming that $\delta$ =1. Note that Oster's (2017, 6) rule of thumb is $\pi$=1.3. The results indicate that for values of $\pi$ < 2.5, the estimate of $\beta_{personalism}$ remains larger than 0.14. This suggests that, given the conservative assumption that the importance of the unobserved variables is the same as the observed variables, the unobserved variables would have to explain more than twice as much of the non-idiosyncratic variation in repression as the observed variables.

[4]Oster's (2017) test does not permit the random effects estimator.

In short, these sensitivity tests suggest that the estimate of personalism is *not* sensitive to conservative and plausible selection on unobservables, easily passing the sensitivity tests proposed by Oster (2017).

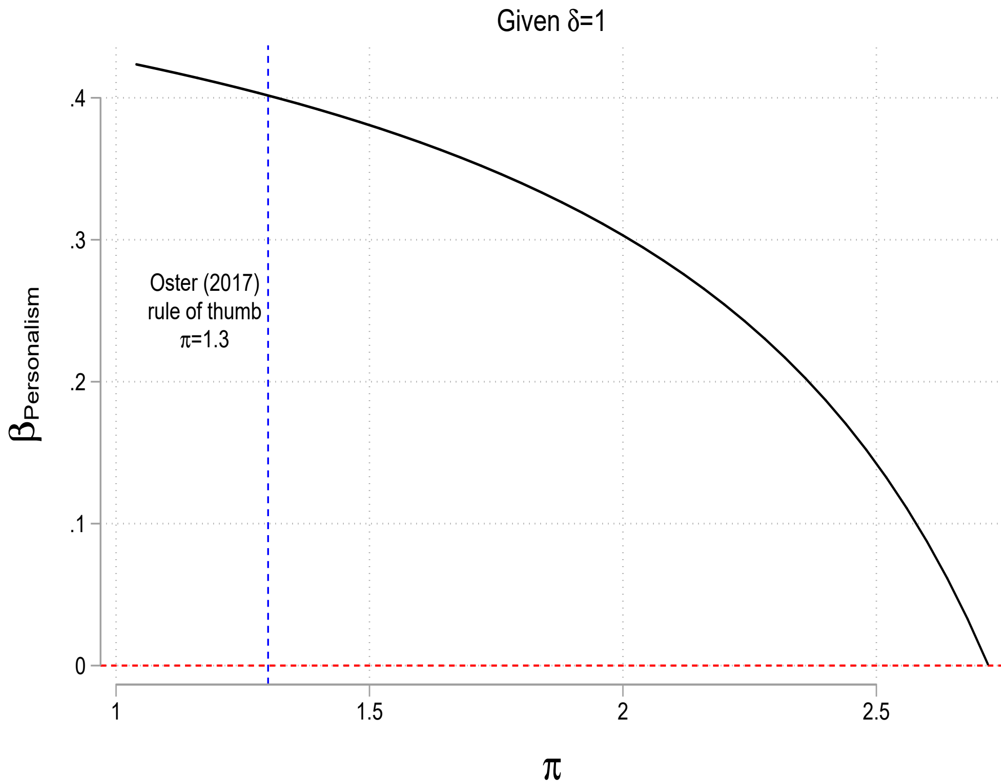


Figure C-1: Sensitivitiy to bias from selection on unobservables.

## Appendix D: Analyzing two sub-components of the personalism index

The personalism index used throughout contains information from eight manifest items identifying when the leader:

- makes access to office dependent on personal loyalty (*Party*);
- creates a new support party after seizing power (*Party*);
- controls appointments to the party executive committee (*Party*);
- makes the party executive committee serve as a rubber stamp for his decisions (*Party*);
- personally controls the security apparatus (*Security*);
- promotes officers loyal to himself or from his support group, or forces officers from other groups to retire (*Security*);
- creates paramilitaries or a new security force loyal to himself (*Security*);
- imprisons or kills officers from other groups without a fair trial (*Security*)

The first four items address the supporting political party, if there is one. [If there is no party, these items are scored at zero.] The latter four items address the leader's relationship with the security apparatus, including the military. In the following analysis, we test two subcomponents of the aggregate personalism index by dividing the items into two groups of four items each, one group for items addressing the party and the other group for item relating to the security appartus. We then create separate subcomponent indices from the four items in each group, producing two subcomponent measures: *Party personalism* and *Security Personalism*. For more work on the latter, see Song (2018). The items are aggregated using a linear combination (Cronbach's alpha) of the items.[5] The standardized indices are rescaled on [0,1].

Figure D-1 reports results from models that test the personalism subcomponents. The first model, in black, reports the results when using the full index (with linear combination of the items rather than the IRT-model index). This result mirrors those reported throughout: increasing personalism is associated with increasing repression. The next model, in blue, reports a test with the Party personalism subcomponent; while positive, the estimate is relatively small and not statstically different from zero. The third model, in green, reports a test with the Security subcomponent; the estimate is positive and significant and similiarly sized to the estimate using the full index. Finally, the fourth model tests both subcomponents at the same time (they are correlated at 0.34). Again, the estimates indicate that personalism in the security apparatus is associated with repression but Party personalism is not.

[5] We use this approach for the full personalism index (i.e. all eight items) in this analysis, for comparability. A IRT 2PL does not converge with four party items.

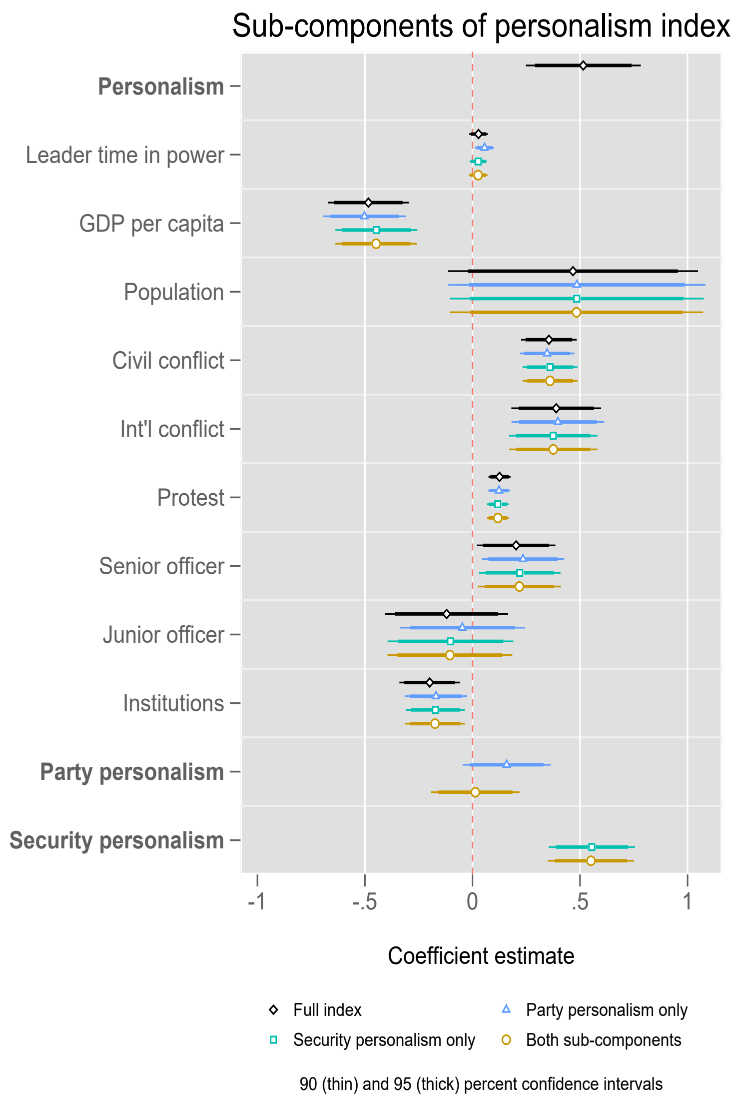


Figure D-1: Party and security sub-components of the personalism index.

# 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.

Chenoweth, Erica and Orion A. Lewis. 2013. "Unpacking nonviolent campaigns: Introducing the NAVCO 2.0 dataset." *Journal of Peace Research* 50(3):415–423.

Chenoweth, Erica, Vito D'Orazio and Joseph Wright. 2014. "A Latent Measure of Political Protest." Paper presented at the International Studies Association's 55th Annual Convention, Toronto (Canada), March 26-29.
**URL:** *http://sites.psu.edu/wright/files/2017/09/CDW-Protest1-239iz5j.pdf*

Crabtree, Charles D and Christopher J Fariss. 2015. "Uncovering patterns among latent variables: human rights and de facto judicial independence." *Research & Politics* 2(3).

Garretón, Manuel Antonio. 1988. "Popular Mobilization and the Military Regime in Chile: The Complexities of the Invisible Transition." Kellogg Institute for International Studies, Working Paper # 103.
**URL:** *https://kellogg.nd.edu/documents/1297*

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

Geddes, Barbara, Joseph Wright and Erica Frantz. 2017. "A Measure of Personalism in Dictatorships." Working paper. October 9, available at `http://sites.psu.edu/dictators/files/2017/10/PersMeasure-1ph2gwp.pdf` (accessed March 14, 2018).

Geddes, Barbara, Joseph Wright and Erica Frantz. 2018. *How Dictatorships Work: Power, Personalization, and Collapse*. New York, NY: 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.

Hill, Daniel W and Zachary M Jones. 2014. "An empirical evaluation of explanations for state repression." *American Political Science Review* 108(3):661–687.

Hyde, Susan D and Nikolay Marinov. 2012. "Which Elections can be Lost?" *Political Analysis* 20(2):191–210.

Jia, Ruixue and Yiqing Xu. 2018. "Rotating to the Top: How Career Tracks Matter in the Chinese Communist Party." Unpublished Manuscript.

Linzer, Drew A and Jeffrey K Staton. 2015. "A global measure of judicial independence, 1948–2012." *Journal of Law and Courts* 3(2):223–256.

Oster, Emily. 2017. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* pp. 1–18.

Pepinsky, Thomas. 2014. “The institutional turn in comparative authoritarianism.” *British Journal of Political Science* 44(3):631–653.

Powell, Jonathan M and Clayton L Thyne. 2011. “Global instances of coups from 1950 to 2010: A new dataset.” *Journal of Peace Research* 48(2):249–259.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.

Shirk, Susan L. 2018. “The Return to Personalistic Rule.” *Journal of Democracy* 29(2):22–36.

Song, Wonjun. 2018. “Dictators, Personalized Security Forces, and Coups.”. Working paper. Available at `https://bit.ly/2ImwbV9` (accessed March 14, 2018).

Urdal, Henrik. 2006. “A clash of generations? Youth bulges and political violence.” *International studies quarterly* 50(3):607–629.

Wilson, Matthew Charles and Joseph Wright. 2017. “Autocratic Legislatures and Expropriation Risk.” *British Journal of Political Science* 47(1):1–17.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. “Ethnic politics and armed conflict: A configurational analysis of a new global data set.” *American Sociological Review* 74(2):316–337.