

## Online Supplementary Appendix

### The Latent Characteristics That Structure Autocratic Rule

November 6, 2019

#### Abstract

Research on autocratic regimes in comparative politics and international relations often uses categorical typologies of autocratic regimes to distinguish among different forms of autocracy. This paper introduces historical data on dozens of features of dictatorships to estimate latent dimensions of autocratic rule. We identify three time-varying dimensions of autocracy that correspond to ideal types proposed in the literature: party dominance, military rule, and personalism. We show that dimensions of autocratic rule are orthogonal to commonly-used measures of democracy-autocracy, and compare these dimensions to existing typologies of autocracies, showing that time-varying information on personalism is unique. We discuss a measurement model of personalism and illustrate the time-varying features of this measure in applied research on conflict initiation.

NOTE: The first paragraph of the original, submitted manuscript contains numerous citations that are omitted in the publication version due to a strict word count limit. It should read as follows:

Research on autocracies in comparative politics and international relations has surged in the past decade. In an effort to understand not just how dictatorships differ from democracies but also to examine how dictatorships differ from one another, this research looks at variation in different forms of non-democratic rule. Studies of international conflict (Peceny, Beer and Sanchez-Terry 2002; Weeks 2012), civil war (Fjelde 2010), international cooperation (Mattes and Rodriguez 2014) and commitments (Leeds, Mattes and Vogel 2009), nuclear proliferation (Way and Weeks 2014), repression (Davenport 2007; Escribà-Folch 2013; Frantz and Kendall-Taylor 2014), international trade (Milner and Kubota 2005; Hankla and Kuthy 2013), terrorism (Aksoy, Carter and Wright 2012; Conrad, Conrad and Young 2014; Wilson and Piazza 2013), foreign investment (Gehlbach and Keefer 2012; Moon 2015), and autocratic survival (Geddes 1999; Smith 2005) demonstrate that variation within the group of countries categorized as non-democracies provides substantial leverage on explaining many important outcomes. This study introduces newly coded historical data to identify the latent dimensions of autocratic rule.

# Contents

<b>1</b>	<b>Appendix A: Data description</b>	<b>1</b>
1.1	Raw data variable definitions . . . . .	1
1.2	Summary Statistics . . . . .	8
<b>2</b>	<b>Appendix B: Additional results for Latent Dimensions</b>	<b>10</b>
2.1	Factor analysis . . . . .	10
2.2	Comparison with democracy measures . . . . .	13
2.3	Comparison with additional measures of autocracies . . . . .	20
<b>3</b>	<b>Appendix C: A measure of personalism from an IRT model</b>	<b>21</b>
3.1	An item-response theory (IRT) approach . . . . .	22
3.1.1	Local independence . . . . .	22
3.1.2	Static model . . . . .	24
3.1.3	IRT estimates . . . . .	25
3.2	Assessing the personalism index . . . . .	28
3.2.1	Comparison with Varieties of Democracy data . . . . .	31
3.2.2	Reliability . . . . .	33
3.2.3	Validity . . . . .	35
3.3	Conclusion . . . . .	37
<b>4</b>	<b>Appendix D: Additional tests for MID initiation</b>	<b>38</b>
4.1	Details of MID initiation results in main text . . . . .	38
4.2	Robustness tests . . . . .	40
4.2.1	Compare autocracies with democracies . . . . .	40
4.2.2	Specifications that exclude the interaction term . . . . .	41
4.2.3	Testing different items in the <i>G-pers</i> measure . . . . .	42
4.2.4	Including early and unstable years of each regime . . . . .	44
4.3	Modeling uncertainty in the latent personalism measure . . . . .	45
4.4	Unpacking personalism and dispute initiation data for China and Libya . . . . .	46
<b>5</b>	<b>Appendix E: Additional tests for regime collapse analysis</b>	<b>48</b>
5.1	Robustness tests . . . . .	48
5.2	Different types of transition outcomes . . . . .	49

# 1 Appendix A: Data description

## 1.1 Raw data variable definitions

These definitions are not the same as the coding rules, which are more detailed. The definitions are only intended to provide readers with a basic understanding of the type of information contained in the raw data. For categorical variables, the items listed in parentheses () fall in the omitted category.

- **leaderrole** (treated as categorical)  
identify how the regime leader achieved office and/or whose support put him in office
  - **\_priordem** prior democracy
  - **\_domparty** dominant party
  - **\_military** military junta
  - **\_insurgency** insurgency
  - **\_hereditary** traditional hereditary succession
  - **\_civsucc** civilian succession
  - **\_interim** interim leader
  - **\_other** other, clerical
  - **\_foreign** foreign imposed
- **seizure** (treated as categorical)  
how did the regime obtain power
  - **\_family** hereditary succession
  - **\_coup** military coup
  - **\_rebel** insurgency/rebels
  - **\_uprising** uprising
  - **\_coup** election
  - **\_succession** authoritarian succession
  - (foreign imposed)
- **ldrexp** (treated as categorical)  
regime leader's most important career experience where he is most likely to have developed his most useful support network
  - **\_highrank** high-ranking military officer
  - **\_lowrank** low-ranking military officer or NCO
  - **\_rebel** leader of armed insurgency that brought regime to power
  - **\_demelect** leader in a prior party organized to run in competitive democratic elections
  - **\_supportparty** position in regime support party and not relative of leader
  - **\_pers\_loyal** the regime leader was chosen by prior regime leaders because of competence or loyalty and NOT officer, party leader or rebel

- `_pers_relative` leader is a close relative of a prior leader of the same regime, who was not himself a hereditary monarch
- `_rulingfamily` member of the traditional ruling family and was chosen in the way that is traditional for the tribe or country in question
- `_other` leader does not fit prior codes
- `supportparty` (binary)
  0. no support party
  1. support party
- `partyleader` (binary)
  0. no support party OR party leader selected by regime leader OR party leader selection influenced by regime leader OR party leader is not regime leader and selection controlled by group that excludes the regime leader
  1. party leader is regime leader or relative of regime leader
- `partyhistory` (treated as categorical)
  - `_noparty` no support party
  - `_postseizure` party created after seizure of power
  - `_priorelection` prior party, created to support autocratic leader election (e.g. Fujimori)
  - `_priornosupport` prior party, never won electoral support
  - `_priorwinsupport` prior party, won support under prior autocracy
  - `_insurgent` insurgent/rebel party
  - `_priordem` prior party elected in a democracy
- `partymins` (ordinal)
  0. no support party
  1. 1/3 or more of cabinet positions go to non-party members
  2. some but fewer than 1/3 of cabinet members are not party members
  3. cabinet ministers (except defense) are party members
- `partymilit` (ordinal)
 

missing if no party or military

  0. military controls party OR no party
  1. no party interference in military
  2. party and military influence each other
  3. party interferes in military but does not impose party structure
  4. party imposes party structure on military
- `partymilit2` (binary)

- 0. not missing on `partymilit`
- 1. no party or military (0 on `partymilit`)
- `partyexcom` (treated as categorical)
  - `pers` regime leader chooses party executive committee
  - `faction` faction that supports the regime leader dominates the party executive committee
  - `oppose` competition for seats on the party executive committee
- `localorgzns` (ordinal)
  - 0. no support party
  - 1. support party has no effect local organizations
  - 2. local-level branch organizations links party militants to citizens
- `excomcivn` (ordinal)
  - 0. no support party
  - 1. party executive committee is 2/3 or more military or retired military
  - 2. party executive committee has military or retired military, but less than 2/3
  - 3. party executive committee is civilian or ex-insurgent
- `multiethnic` (binary)
  - 0. no support party OR monoethnic
  - 1. party leadership is multi-ethnic/region/religious
- `monoethnic` (binary)
  - 0. no support party OR multiethnic
  - 1. party leadership is dominated by people from particular ethnic/region/religious
- `heirparty` (binary)
  - 0. not 1
  - 1. heir is high party official but not close relative of the leader
- `heirfamily` (binary)
  - 0. not 1
  - 1. heir is same family as a leader before him within the same regime
- `heirciv` (ordinal)
  - 0. military succession
  - 1. leader from insurgency

- 2. civilian succession
- **heirclan** (binary)
  - 0. regime leader not from same clan etc or ethnicity/clan/tribe not politically relevant
  - 1. regime leader (or heir apparent) from same clan, tribe, or ethnic group as successor
- **legcompetn** (ordinal)
  - 0. no legislature
  - 1. appointed by regime leader
  - 2. indirect selection of legislative body by another body
  - 3. all seats from uncontested elections
  - 4. only front groups and ruling party members
  - 5. all seats from ruling front/party, but competitive multi-candidate elections
  - 6. only independents seated in opposition
  - 7. some opposition seats from elections but less than 25% (includes independents)
  - 8. 25% or more opposition seats from elections (includes independents)
- **leaderciv** (binary)
  - 0. leader was NOT civilian before being in power
  - 1. leader was civilian before being in power
- **leadermil** (binary)
  - 0. leader was NOT member of the military before assuming power
  - 1. leader was member of the military before assuming power
- **leaderrebel** (binary)
  - 0. leader was NOT member of an insurgency before assuming power
  - 1. leader was member of an insurgency before assuming power
- **cabciv** (ordinal)
  - 0. most important cabinet positions held by military OR regime leader
  - 1. cabinet is civilians or insurgents, but some military in positions other than defense
  - 2. civilian cabinet (except defense)
- **cabmil** (ordinal)
  - 0. most positions (except defense) held by civilians OR regime leader
  - 1. cabinet is civilians or insurgents, but some military in positions other than defense
  - 2. most important cabinet positions held by military

- `militrank` (ordinal)
  0. leader was not a (retired) member of the military; has honorific military title; or was member of an insurgency
  1. leader was rank below major
  2. leader was a colonel in a military that includes generals
  3. leader was a colonel in a military that did not include generals
  4. leader was general, admiral, or other highest ranking office
- `milmerit_pers` (ordinal)
  0. regime leader does not use loyalty in promotion AND no widespread forced retirement OR no military
  1. promotions of top officers loyal to the regime leader or from his group
  2. regime leader promotes officers loyal to himself or from his ethnic, tribal, regional, or religious group OR widespread forced retirements
- `milmerit_mil` (ordinal)
  0. officer promotion based on personal loyalty to leader OR widespread forced retirement OR no military
  1. promotions of top officers loyal to the regime leader or from his group
  2. regime leader does not promote officers loyal to himself or from his ethnic, tribal, regional, or religious group AND no widespread forced retirement
- `milconsult` (binary)
  0. no consultative body; regime leader not from the military
  1. consultative body in which the heads of service branches are represented; or if country specialists describe some other routinized method of consultation
- `milnotrial` (ordinal) regime leader imprison/kill officers or officers from other groups without fair trial
  0. regime leader does NOT kill/imprison out-group officers OR no military OR foreign officers staff military
  1. regime leader imprisons/kills officers from other groups without fair trial
- `militparty` (treated as categorical)
 

has regime leader created his own support party

  - `_noparty` regime leader's supporters are not organized in a party
  - `_ally` regime leader allies with a pre-existing party
  - `_newparty` regime leader or a close ally creates a party to support the regime after his accession to office

- `_priorparty` regime leader’s supporters were organized into a party prior to his accession to power, and that party now supports the regime
- `milethnic` (treated as categorical)
  - officers from more than one ethnic, religious, or regional group
  - `_dom` high ranking officers come from most of the larger, politically salient ethnic, religious, and regional groups OR not salient cleavage
  - `_hetero` salient cleavage AND one or a few regions, ethnicities, or religions are overrepresented in officer corps AND officers include some high ranking members from different backgrounds
  - `_homo` salient cleavage AND nearly all high ranking officers come from one or a few regions or groups
- `nomilitary` (binary)
  - 0. military present AND coded as one of three types of `milethnic`
  - 1. no military or officers are mostly foreign
- `ldrrotation` (binary)
  - 0. no rotation procedure; or regime leader is not from the military
  - 1. procedure for regular succession or rotation of the executive among military officers (including rigged elections)
- `electldr` (treat as categorical)
  - `_notelect` not elected
  - `_priordict` elected in prior dictatorship
  - `_1candidate` elected in one candidate election
  - `_1party` elected in one party election
  - `_multileg` selected by legislature elected in multiparty elections
  - `_multiexec` elected in multiparty executive elections
  - `_priordem` elected in prior democracy
- `legnoms` (treat as categorical)
  - 0. (no opposition in legislative elections)
  - 1. `_indirect` legislature selected by indirect election from lower body; comprised of local/tribal notables; or selected by regime insiders as societal representatives
  - 2. `_veto` opposition allowed to contest but regime holds veto power of candidate selection
  - 3. `_noveto` opposition or independents allowed to contest; and ruling party candidate selection influenced by local party leaders or faction members
  - 4. `_priordem` legislature chosen in prior democratic regime or competitive pre-independence election

- **plebiscite** (binary)
  - Has the regime leader held a plebiscite to legitimize or consolidate his rule
    0. no plebiscites on the regime leader's occupancy of the executive or the continuation of the regime have been held
      1. one or more plebiscites have been held
- **partyrbrstmp** (binary)
  0. party executive committee as a rubberstamp; no party executive committee; no support party
    1. party executive committee has some policy independence from the regime leader
- **officepers** (binary)
  0. regime leader does not have discretion over appointments to high office or appoints relatives to these positions
    1. regime leader has discretion over appointments to high office or appoints relatives to these positions
- **leaderreatvls** (binary)
  0. regime leader does not have discretion over appointments to high office or appoints relatives to these positions
    1. regime leader has discretion over appointments to high office or appoints relatives to these positions
- **paramil** (treat as categorical)
  0. **\_fightrebel** paramilitary forces created to fight civil war on regime's side
    1. **\_party** party militia or paramilitary organized by dominant party
    2. **\_pers** regime leader creates paramilitary forces, a president's guard, or new security forces apparently loyal to himself
- **secty\_app** (treated as categorical)
  - **\_party** security apparatus controlled by dominant party
  - **\_pers** security apparatus controlled personally by regime leader

## 1.2 Summary Statistics

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
partyrbrstmp	0.303	0.46	0	1	4591
militrank	1.068	1.625	0	4	4591
ldrrotation	0.032	0.175	0	1	4591
milconsult	0.134	0.341	0	1	4591
milmerit_mil	0.732	0.767	0	2	4591
milmerit_pers	1.186	0.791	0	2	4591
milnotrial	0.365	0.482	0	1	4591
plebiscite	0.222	0.415	0	1	4591
heirclan	0.373	0.484	0	1	4591
officepers	0.644	0.479	0	1	4591
ParamilPers	0.354	0.478	0	1	4591
ParamilParty	0.184	0.388	0	1	4591
ParamilFReb	0.088	0.283	0	1	4591
supportparty	0.731	0.443	0	1	4591
partyleader	0.588	0.492	0	1	4591
localorgzns	1.345	0.874	0	2	4591
partymins	1.783	1.326	0	3	4591
excomcivn	1.543	1.206	0	3	4591
multiethnic	0.504	0.5	0	1	4591
monoethnic	0.225	0.418	0	1	4591
heirparty	0.429	0.495	0	1	4591
heirfamily	0.393	0.488	0	1	4591
legcompetn	4.166	2.971	0	8	4591
leaderrelatvs	0.504	0.5	0	1	4591
leaderciv	0.549	0.498	0	1	4591
leadermil	0.346	0.476	0	1	4591
leaderrebel	0.105	0.307	0	1	4591
heirciv	0.92	0.983	0	2	4591
cabciv	1.322	0.729	0	2	4591
cabmil	0.598	0.688	0	2	4591
partymilit	1.083	1.566	0	4	4591
ldrPriorD	0.09	0.287	0	1	4591
ldrParty	0.237	0.425	0	1	4591
ldrMil	0.287	0.453	0	1	4591
ldrRebel	0.103	0.304	0	1	4591
ldrCiv	0.017	0.128	0	1	4591
ldrOther	0.059	0.236	0	1	4591
ldrForgn	0.092	0.289	0	1	4591
ldrHered	0.114	0.318	0	1	4591
SeizCoup	0.303	0.46	0	1	4591
SeizRebel	0.221	0.415	0	1	4591
SeizUpris	0.037	0.19	0	1	4591
SeizElec	0.132	0.339	0	1	4591
SeizSucc	0.043	0.202	0	1	4591

*Continued on next page...*

... table 1 continued

Variable	Mean	Std. Dev.	Min.	Max.	N
SeizFam	0.056	0.23	0	1	4591
partymilit2	0.551	0.497	0	1	4591
MilPartyNew	0.117	0.321	0	1	4591
PartyhNoWin	0.077	0.266	0	1	4591
PartyhWin	0.019	0.138	0	1	4591
PartyhReb	0.16	0.367	0	1	4591
PartyhPriorDem	0.239	0.426	0	1	4591
PartyhNoparty	0.269	0.443	0	1	4591
PartyhPost	0.221	0.415	0	1	4591
PartyhElec	0.016	0.125	0	1	4591
MilPartyAlly	0.032	0.177	0	1	4591
MilPartyNo	0.112	0.315	0	1	4591
MilPartyPrior	0.076	0.265	0	1	4591
nomilitary	0.041	0.198	0	1	4591
milethnic_inclusive	0.458	0.498	0	1	4591
milethnic_hetero	0.34	0.474	0	1	4591
milethnic_homo	0.162	0.368	0	1	4591
sectyapp_party	0.155	0.362	0	1	4591
sectyapp_pers	0.596	0.491	0	1	4591
ElecldrNot	0.362	0.481	0	1	4591
ElecldrPrDict	0.009	0.094	0	1	4591
ElecldrPrDem	0.021	0.144	0	1	4591
ElecldrNot	0.362	0.481	0	1	4591
Elecldr1C	0.217	0.412	0	1	4591
Elecldr1F	0.038	0.192	0	1	4591
ElecldrMLeg	0.055	0.229	0	1	4591
ElecldrMExec	0.174	0.379	0	1	4591
legnoms_indirect	0.088	0.283	0	1	4591
legnoms_veto	0.374	0.484	0	1	4591
legnoms_noveto	0.114	0.318	0	1	4591
legnoms_priordem	0.019	0.136	0	1	4591
LdrexHighR	0.233	0.423	0	1	4591
LdrexLowR	0.093	0.291	0	1	4591
LdrexRebel	0.11	0.313	0	1	4591
LdrexDemEl	0.134	0.341	0	1	4591
LdrexParty	0.232	0.422	0	1	4591
LdrexLoyal	0.019	0.137	0	1	4591
LdrexReltv	0.027	0.162	0	1	4591
LdrexRulFam	0.128	0.334	0	1	4591
LdrexOther	0.023	0.15	0	1	4591
partyexcom_pers	0.318	0.466	0	1	4591
partyexcom_faction	0.22	0.414	0	1	4591
partyexcom_oppose	0.127	0.333	0	1	4591
createparty	0.162	0.368	0	1	4591

## 2 Appendix B: Additional results for Latent Dimensions

### 2.1 Factor analysis

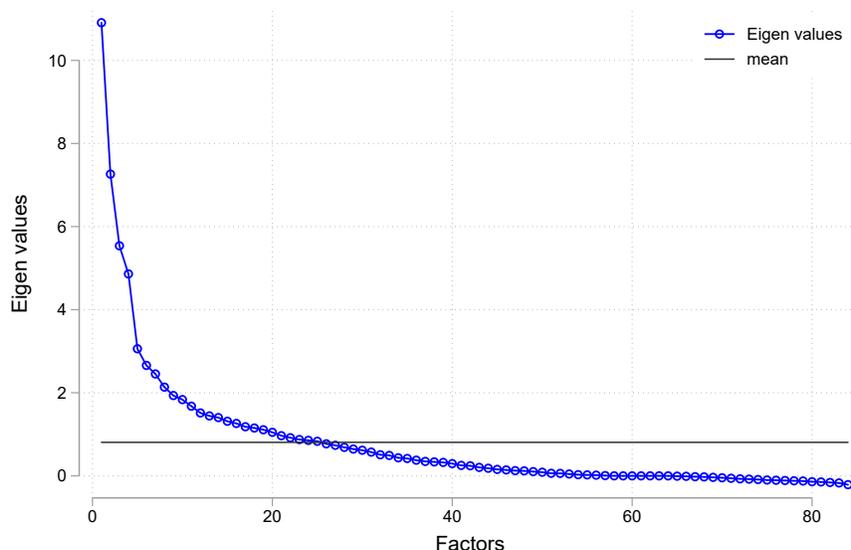


Figure B-1: Eigenvalues. The eigen values are shown for the first 8 dimensions

The first six factors contain 50 percent of the information for 85 items. The first 3 factors, which we use in the analysis, contain 35 percent of the total variation. The items that load most strongly on the fourth factor, which we drop from the analysis, are: regime seizure in a rebellion; leader experience is as leader of a rebellion; leader is a rebel leader; and the support party existed prior to regime seizure as a rebel party.

The plots in Figure B-2 show factors in a three-dimensional space derived from constructing the first three dimensions. Each panel shows two dimensions plotted against each other. Each circle in the plots represents a variable employed in the factor analysis. The distance of each circle from the center shows how much each variable contributes to the two dimensions represented along each axis – i.e. how much weight each variable contributes to the factor loading. Variables that align in opposite directions (180 degrees) contain similar information but point in different directions because of the (necessarily) arbitrary ordering of the information contained in the variable.

The upper left plot shows the first two dimensions. An interpretation of the first dimension, along the horizontal axis, is that it measures the extent to which the political party supporting the regime has power. While difficult to distinguish visually, the cluster of variables on the right of the plot contain information on this concept: whether a support party exists (`supportparty`); whether the support party has local-level branch organizations that link party militants to citizens (`localorgs`); and whether all the cabinet appointees, except defense, are members of the ruling party (`partymins`).

The second dimension, displayed on the vertical axis in the upper left plot (and on the horizontal axis in the lower left plot), measures the extent to which the military – as an institution – has power. The variables that contribute substantial information to this dimension relate to military power,

for example whether the military selected the regime leader (**LdrMil**); whether the regime leader was a high ranking military officer prior to assuming power (**militrnk**); whether there is a routine mechanism for the leader to consult the military in policy decisions (**milconsult**); and the extent to which military officers comprise the cabinet (**cabmil**). The third dimension, shown on the vertical axis in the upper right plot and the vertical axis on the lower plot, measures the level of personalist power for the regime leader. It captures concepts such as whether the regime leader rather than the ruling party controls the security apparatus (**sectyapp\_pers** and **sectyapp\_party**) and whether the dictator personally controls appointment to high office (**officepers**).

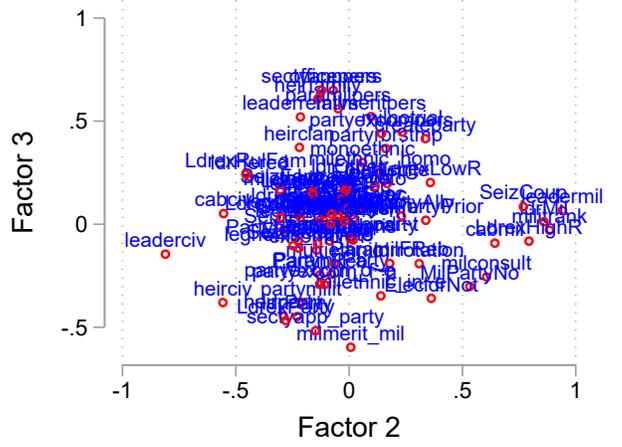
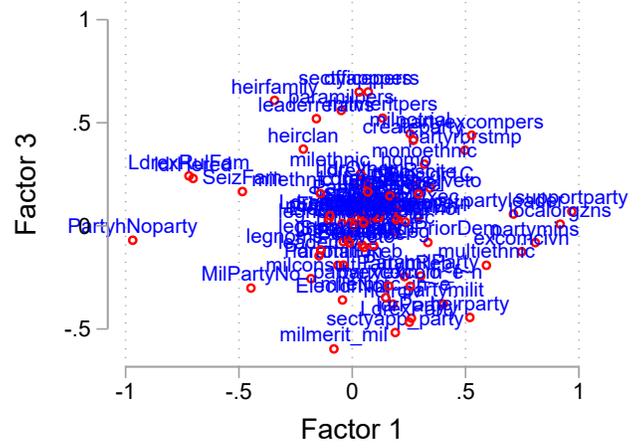
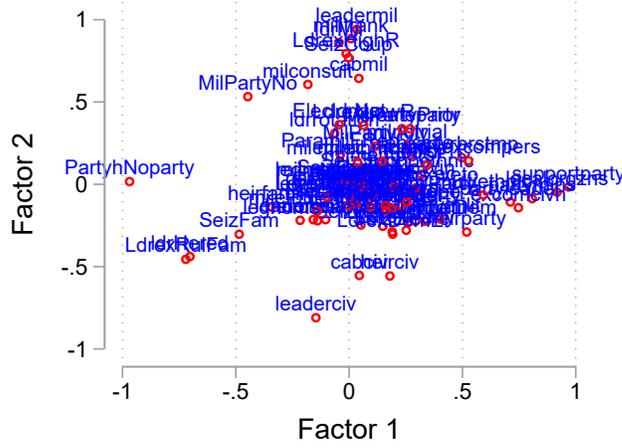


Figure B-2: Three features of autocratic rule.

## 2.2 Comparison with democracy measures

Table B-2 shows the correlation between the ‘within’ variation for the personalist variable and the ‘within’ variation for measures of various democratic concepts, from the Varieties of Democracy (vdem) project and the Polity (polity) project.<sup>1</sup> The first four row variables, are distinct dimensions of democracy, as defined by the VDem project: liberal, participatory, deliberative, and egalitarian. The next two row variables measure different types of government accountability, vertical and horizontal. The next three row variables, also from VDem, measure the negative power of the head of state to control other members of the government: dissolve the legislature, veto legislation, and dismiss ministers. Importantly, these negative powers are measured in practice, not by the formal rules of the political system.

The first column lists the variable name, while second lists the source for the extant variable. The third column shows the ratio of ‘within’ to total variance for the variable in each row. The variable with the highest ratio of ‘within’ variation is *vertical* accountability from VDem: 0.417. The final column shows the correlation between the within variation in the row variable and the within variation in the personalism measure. From this last column, we can see that the within variation in the personalism measure (dimension 3 in the factor analysis) is not highly correlated with the within information in any of the measures of democratic concepts.

The main point of this table is to show that personalism is not correlated with any sub-components of democracy as measured by these variables. As important, this analysis isolates the ‘within’ variation, to highlight how the within variation in the personalism measure is unique and thus likely to capture unmodeled variation in applied research using fixed effects designs.

variable	source	$\frac{\textit{within}}{\textit{within}+\textit{between}}$ variation ratio	within $Cov(\textit{variable}, \textit{personalist})$
liberal	vdem	0.274	-0.174
partipatory	vdem	0.283	-0.032
deliberative	vdem	0.323	-0.129
egalitarian	vdem	0.197	0.086
vertical	vdem	0.417	0.006
horizontal	vdem	0.286	-0.157
dissolve legis	vdem	0.243	0.183
veto legis	vdem	0.237	0.163
dismiss minister	vdem	0.224	0.157
xrreg	polity	0.278	0.035
xrcomp	polity	0.246	-0.08 9
xropen	polity	0.231	0.033
xconst	polity	0.292	-0.108
parreg	polity	0.367	0.04
parcomp	polity	0.361	-0.041
polcomp	polity	0.368	-0.054

Table B-2: Within correlation between Personalist dimension and democracy components

Figure B-3 shows the Spearman correlations between the latent dimensions (Party, Military, and Personal) as well as two extant democracy measures (Polity and Polyarchy) and various subcom-

<sup>1</sup>The ‘within’ variation is the information that changes over time within the regime-case unit.

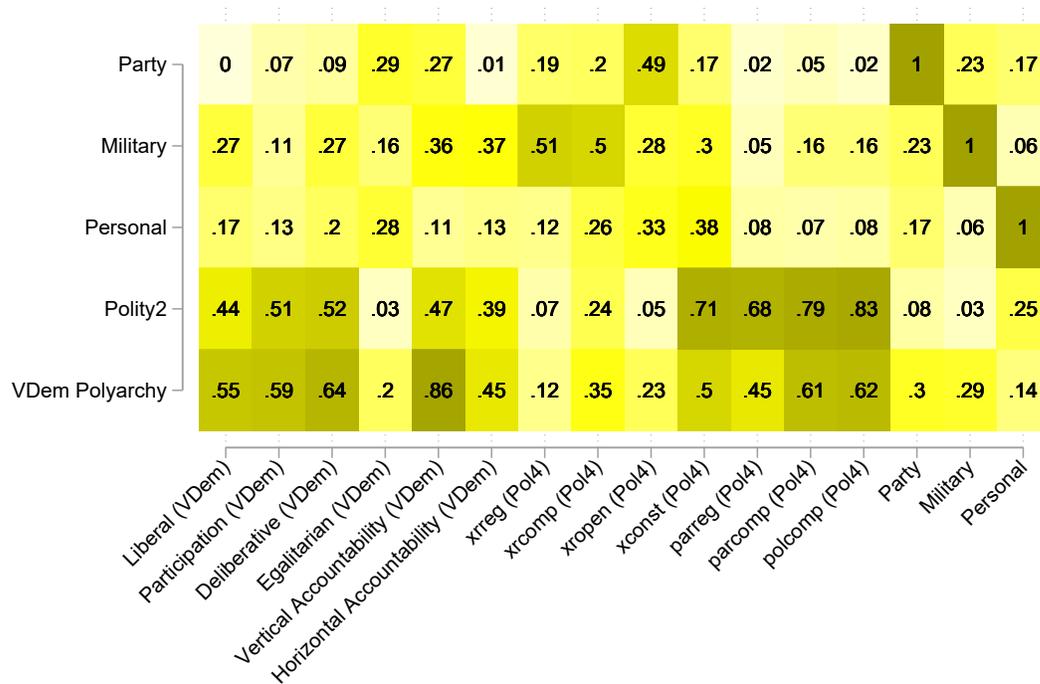


Figure B-3: A correlation matrix of our three latent dimensions of autocratic structure, and several components of democracy. Visually, darker cells show higher correlation. Democracy scores largely group together as a high correlation block, while our latent dimensions are not in that block. None of these measures correlate highly with the third Personalist latent dimension

ponents of democracy, as measured by variables drawn from VDem and Polity. First, note the dark block of correlations in the lower right, showing that subcomponents of Polity (xconst, parreg, parcomp, and polcomp) are highly correlated with the Polity2 variable as well as VDem’s Polyarchy. None of these Polity components are strongly correlated with the latent dimensions of autocracy. Second, note that the various measures of democracy from VDem (liberal, participation, deliberative, egalitarian, and accountability) are correlated with the latent autocracy dimensions. Third, three components of Polity – xrreg, xrcomp, and xropen – are correlated with a dimension of autocracy. The first two, xrreg and xrcomp, are strongly negatively correlated with the military dimension, while xropen is positively correlated with the party dimension. These relatively high correlations result from the fact that some autocracies with high party autonomy have elected leaders (e.g. PRI in Mexico); and many autocracies with more autonomous militaries have leaders who come to power in coups, which reflect low competitiveness and regulation of leadership selection. Finally, the personalism dimension of autocracy is not highly correlated with any of the democracy measures.

In Figure B-4 we present simple summary statistics of our autocratic measures of the latent dimension, across the range of Polity scores. We only include data for autocratic countries; the bar chart on the bottom of the graph shows the relative number of observations in our data at each level of Polity. Most of the data is below -5 on this index, and very little is above a score of

0. At the very lowest bound of the polity score, our measures are very different than elsewhere in the data. Figures B-5, B-6, B-7 show that the regime-years at polity of -10 are tightly clustered in our own latent space. However, across the rest of the range where we have frequent data, from -9 to 4, there is no systematic relationship between the means of either of our latent dimensions and the associated polity score. In the rare cases in the data where the polity score is positive, the first dimension mean is distinct, and greater than that in the rest of the data. Generally then, the means of our measures of autocratic structure are uninformative and unrelated to the level of democratization, except at the bounds of the data.

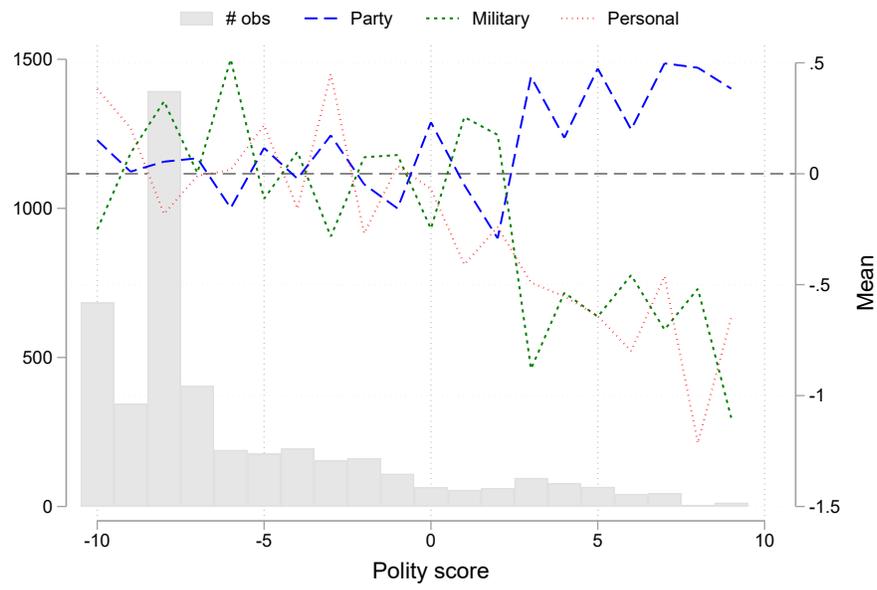


Figure B-4: Means of latent dimensions, across levels of democracy. Outside the bounds of the data (-9 to 3), the mean positions appear unrelated to the Polity score.

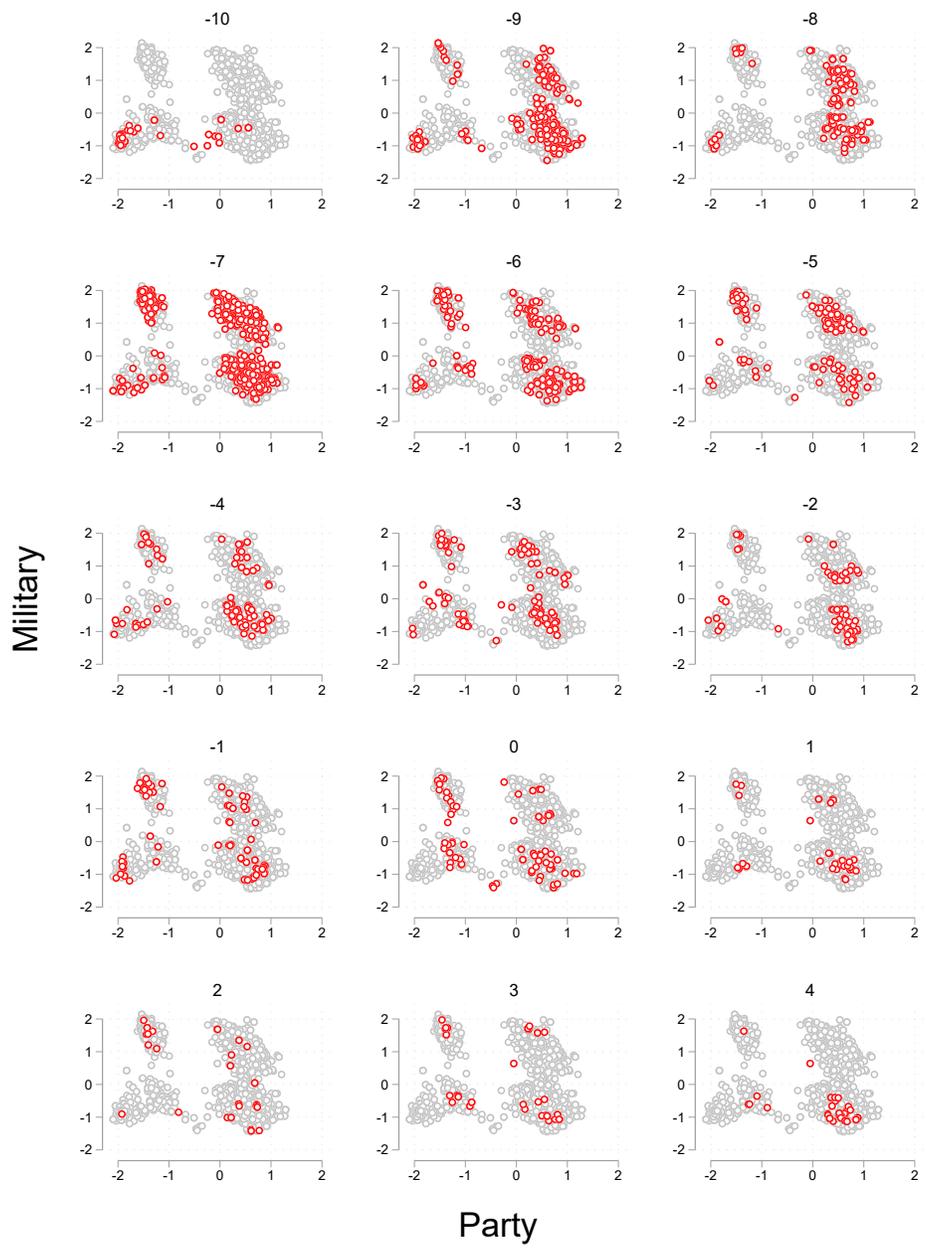


Figure B-5: First (party) and second (military) latent dimensions at each level of the Polity score.

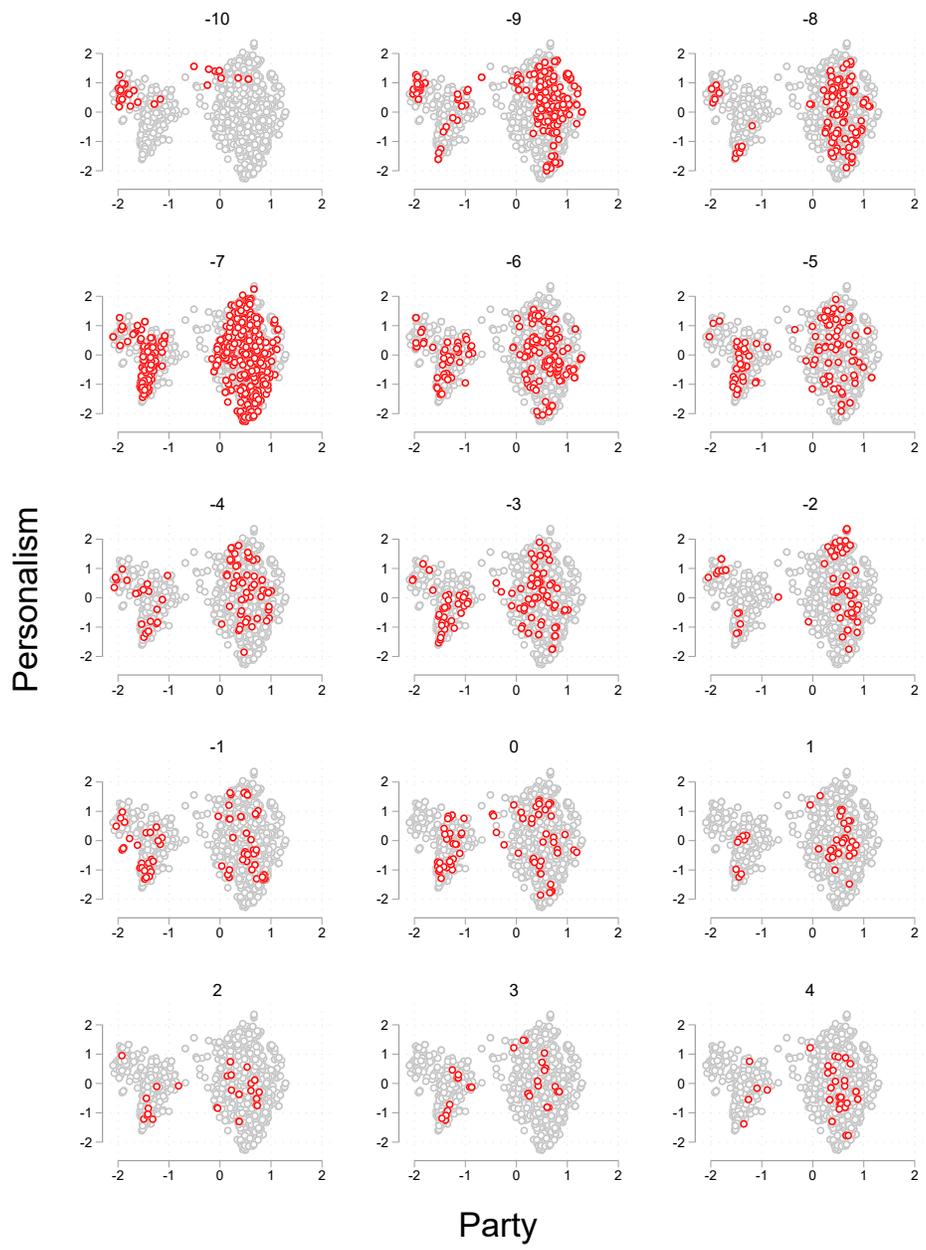


Figure B-6: First (party) and third (personalism) latent dimensions at each level of the Polity score.

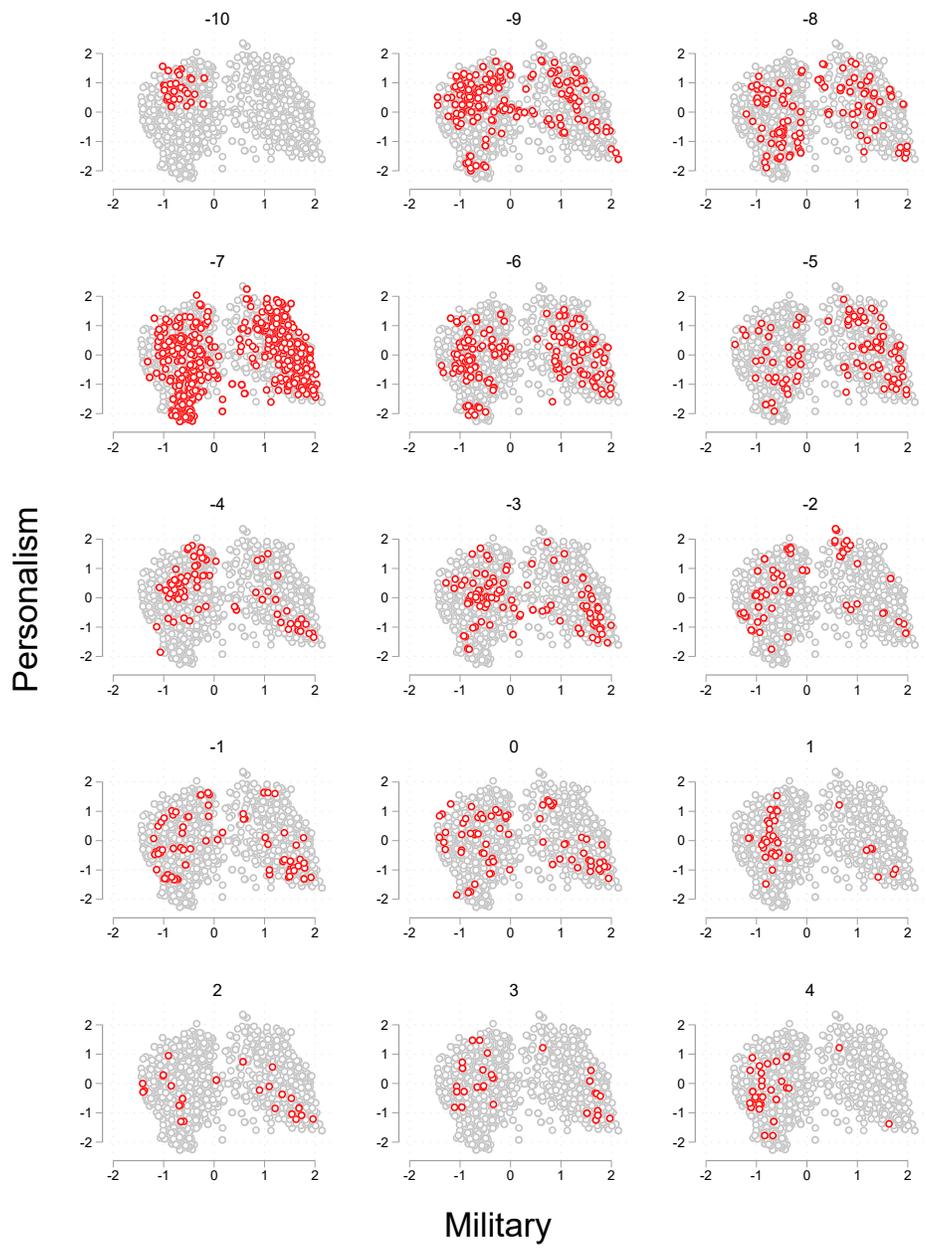


Figure B-7: Second (military) and third (personalism) latent dimensions at each level of the Polity score.

## 2.3 Comparison with additional measures of autocracies

Table B-3: Pairwise bivariate correlations

Variables	dimension 1	dimension 2	dimension 3
Number of parties (Gandhi)	0.328	-0.03	-0.089
Number of parties (H&T)	0.199	0.033	-0.065
Number of parties (DPI)	0.419	-0.283	-0.047
Number of parties (Svolik)	0.375	-0.001	-0.108
One or more parties (Gandhi)	0.624	-0.11	-0.144
One or more parties (H&T)	0.63	-0.292	-0.119
One or more parties (DPI)	0.517	-0.347	-0.049
One of more parties (Svolik)	0.662	-0.054	-0.178
Leader associated w. party (Svolik)	0.724	-0.21	-0.03
Institutions (Gandhi)	0.461	-0.242	-0.059
Legislative competitive index (Svolik)	0.301	-0.263	0.03
Legislative competitive index (DPI)	0.389	-0.28	-0.021
Military regime (Gandhi)	-0.039	0.854	0.186
Military regime (H&T)	-0.176	0.735	0.119
Military regime (DPI)	-0.028	0.799	0.284
Military regime (Svolik)	-0.095	0.846	0.176
Corporate military (Svolik)	-0.166	0.554	0.027
Personal military (Svolik)	0.042	0.5	0.216
Indirect military (Svolik)	-0.028	0.13	-0.066

### 3 Appendix C: A measure of personalism from an IRT model

In this Appendix, we describe an item response theory (IRT) approach for modeling latent personalism. We begin by examining the factor loading plot from the exploratory factor analysis (see Figure 1 in the main text) to identify items that contribute substantial information to a personalism dimension of autocracy. Table C-1 lists the 11 items with a factor loading of 0.35 or more on the third, personalism, dimension. Each of these 11 items is a binary indicator, so the mean value represents the share of observations for which the indicator equals 1.<sup>1</sup>

We treat these items as behavioral manifestations of the autocratic leader’s power relative to other elites in the regime. Items related to the security apparatus, for example, identify when the leader takes personal control over the security apparatus (**SectyappPers**), purges senior officers (**milnotrial**), creates a personally loyal paramilitary (**ParamilPers**), and promotes loyal officers in the military (**MilmeritPers**). These items do not include information about how the leader seized power or a personal attribute about the leader’s prior experience (e.g. in the military or supporting political party).

Both the strength of the factor loadings and the frequency with which items are equal to 1 vary considerably across items. The two items that load most strongly in the factor analysis model, **SectappPers** and **officepers** are also the most frequently observed. However, there are also items, such as **ParamilPers**, that are less frequently observed but still load strongly. Still other items, such as **createparty**, are not observed in most cases (16 percent of observations). Because the mean values and the factor loadings differ substantially across items, the items are likely have distinct difficulty (related to the mean) and discrimination (related to the factor loading) estimates in an IRT model.

Table C-1: Item summary

<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Factor Loading</b>
SectyappPers	0.596	0.491	0.649
officepers	0.644	0.479	0.649
heirfamily	0.393	0.488	0.609
ParamilPers	0.354	0.478	0.560
leaderrelatvs	0.504	0.5	0.520
MilmeritPers	0.423	0.494	0.522
milnotrial	0.365	0.482	0.447
PartyExcomPers	0.318	0.466	0.440
createparty	0.162	0.368	0.415
heirclan	0.373	0.484	0.372
partyrbstmp	0.303	0.46	0.367
Observations		4591	

In what follows we examine model specifications with all 11 items, with 10 items (dropping the

<sup>1</sup>**MilmeritPers** is collapsed from an ordered variable with three values (0, 1, and 2). Replication files contain a mixed model that uses a graded response function for this ordered variable if users are interested in using the ordered information in this variable.

item with the lowest discrimination parameter estimate, `heirclan`), and with 8 items (dropping `heirclan`, `heirfamily`, and `leaderrelative`). Before walking through estimates for various model specifications, we first discuss the IRT modeling approach and potential violations of the model assumptions when using this data.

### 3.1 An item-response theory (IRT) approach

We employ a 2-parameter logistic (IRT-2PL) model, where  $i$  indexes regimes,  $t$  indexes calendar years, and  $j$  indexes the 11, 10, or 8 items that we posit can be treated as observable indicators (with measurement error) of personalism.

$$Pr(y_{j,i,t} = 1 | \text{personalism}_{i,t}) = \text{logit}(\delta_j + \beta_j \text{personalism}_{i,t}) \quad (1)$$

In this equation,  $\delta_j$  is the difficulty parameter;  $\beta_j$  is the discrimination parameter for item  $j$ ; and the logit function is a logistic transformation of the data. The purpose of the equation is to estimate  $\text{personalism}_{i,t}$ , which is the estimated degree of personalism for each regime-year. The difficulty parameter ( $\delta_j$ ) reflects the extent to which regimes, on average, are observed to have political events that correspond with one of the items, while the discrimination parameter ( $\beta_j$ ) reflects the extent to which one item predicts another item.

#### 3.1.1 Local independence

One assumption of the IRT model is local independence of the item responses, meaning response to one item does not influence response to another item conditional on the value of the latent trait. The data-collection process in this project may lead to violations of the local independence assumption when observed (historical) political events serve as manifestations for more than one item. For example, historical evidence of personal control over the security apparatus `security apparatus` may include (among other types of events that constitute evidence for this item): the leader creating new security force under his control or the leader appointing a family member as the head of a security force. Evidence of the former event (creating a new security force) may also be used as evidence for the `paramilitary` item. And appointing a family relative as head of a security force can also serve as an indication for the `high office` item. By construction, these items may be related because the same historical event serves as the basis for coding a manifest indicator for more than one item (variable). If one views items as multiple (flawed) attempts to measure an underlying personalism trait rather than as independent tests of a latent trait, then this point is not an issue.

Another possible source of item dependence is structural missingness in the data. For example, regimes that do not have a supporting political party will not have created a new political party (`create new party`), will not have a rubber stamp party (`rubber stamp party`), and the leader will not control appointments to the party executive committee (`party exec committee`). Thus responses on these items will be related (they are coded as 0 for each of these items) due to an unmodeled regime feature, in this case, lack of a supporting political party. This fact cannot be avoided because the real world of dictatorships creates this type of structural missingness.<sup>2</sup>

---

<sup>2</sup>An additive index of personalism cannot address structural missingness either.

Violations of local independence can lead to artificially inflated scores that may compromise reliability. One piece of evidence that might indicate a violation of local independence is the presence of a second dimension in the items used to estimate the IRT model. If, conditional on the estimate of the latent (personalism) trait, the items are related to one another via a second, unmodeled dimension, this would constitute a violation of the local independence assumption.

Our goal in this appendix is to estimate an IRT model that produces a comprehensive measure of personalism; this may not lead to the construction of the most reliable measure if the information in the items are not independent, conditional on the estimate of the latent trait. To understand why there may be more than one dimension even in this reduced number of items (11), we note that these items incorporate information about two underlying aspects of personalism: personalism related to the security sector, including the military, and personalism related to the support party. Both aspects of personalism might be modeled separately in future research; in doing, so future research will likely produce measures that more closely meet the local independence assumption.<sup>3</sup>

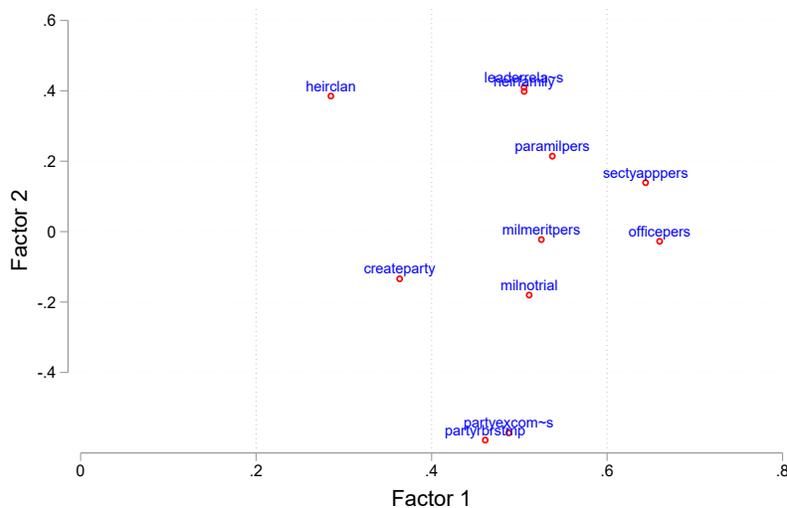


Figure C-1: *Factor loadings for 2 dimensions of personalism.*

To explore the possibility of a second dimension in the information contained among these 11 items listed in Table C-1, we estimate a factor model. The eigenvalue for the first factor is 2.85; that for the second is 1.27; and that for the third is 0.41. The eigenvalues suggest the possibility of two dimensions among these 11 items. Figure C-1 shows the factor loadings for the first two dimensions from a factor model that uses the 11 items. The first factor, depicted along the horizontal axis, shows that all items load in the same direction on this dimension, though with varying strength. The item `heirclan` has the weakest loading, while `officepers` has the strongest. There is also substantial variation among the items along the vertical axis, which depicts the second factor. The items `heirclan`, `heirfamily`, and `leaderrelatvs` score strongly in one direction on this dimension, while `PartyExcomPers` and `partyrbrstamp` score in the other direction on this second dimension.

<sup>3</sup>See “Author” (2017b) for an application of this data to the North Korean case. They trace the sequence of personalism to show that the first regime leader, Kim Il-Sung, consolidated power of the security sector prior to personalizing power over the Korean Workers Party.

Those three strongly scoring items capture the presence of informal institutions relating to the leader’s family and clan predominate, while the the latter two items indicate the presence of a support party, though one that is unlikely to constrain the leader’s behavior. Thus the second dimension may be measuring the level formal institutionalization in the regime. This plot suggests one item in particular, `heirclan`, that we might considered dropping for the IRT model because it loads most weakly on the first dimension and loads strongly on the second, which means it likely contributes to a violation of the independence assumption.

One way to address potential violations of local independence is to drop items that are related. We do this iteratively below. Another way to address local dependence is to combine related indicators into an ordered indicator and use a graded response IRT model. We show in the accompanying code that this approach, while reducing between-item dependence, yields almost identical information as the 8-item latent measure we introduce below.<sup>4</sup> We leave future research to build a personalism measure that directly models the structural missingness, particularly with respect to the presence (absence) of a supporting political party, with the aim of addressing dependence across items.

### 3.1.2 Static model

The IRT model implemented here is static: the estimates of the latent trait are determined only by the information conveyed in the items for that country year, without accounting for possible temporal dependence in the data. However, the estimates of the latent trait for the same leader over multiple years in power are unlikely to be independent because the data were coded with instructions that imply *dependence* across some item responses (over time) for the same individual leader. For example, instructions for coding the `military promotion` item state that coders should “use the same codes for a leader’s entire time in office unless observers specifically mention changes in the leader’s promotion strategy.” This coding rule likely creates dependence over time in item responses for the same regime leader. This coding procedure assumes that once a leader has grabbed sufficient ‘power’ from other elites for coders to observe a manifestation of this indicator, this ‘power’ remains as a latent capacity held by the leader, despite manifestations not being observed continuously thereafter for individual leaders’ duration as de facto regime leader. The data therefore mark observed manifestations of an increase in the leader’s power, not decreases in this power – unless explicitly observed from reading case studies of the regime.

One way to address dependence over time for individual leaders is via a dynamic IRT model that uses the prior estimate of the latent trait ( $\theta_{i,t-1}$ ) for a particular leader  $i$  as the center of the distribution from which to draw the prior of the latent trait estimate ( $\theta_{i,t}$ ) for  $i$  in subsequent period  $t$ ; i.e.  $\theta_{i,t} \sim N(\theta_{i,t-1}, k)$  where  $k$  is fixed.<sup>5</sup>

One rationale for estimating dynamic models is to enhance efficiency so the model will yield smaller standard errors around the estimate of the latent trait. An alternative rationale for using dynamic models is to account for missing data. However, there is no missing data for any of the manifest items, so a dynamic model does not provide substantial additional information from the lagged estimate to help estimate the latent trait for observations for which some items are missing.

---

<sup>4</sup>The Spearman correlation is 0.97, while the Spearman correlation for the *within-leader* variation is 0.93.

<sup>5</sup>For example, Martin and Quinn (2002) set  $k = 1$ , while Gandhi and Sumner (2017) set  $k = 0.25$ . Kenwick (2017) draws  $k$  from a uniform distribution over (0,1).

When weighing the benefits of a dynamic IRT in this application, it is important to note that some of the cross-section units (i.e. distinct autocratic leaders) over which we would want to relax the independence assumption are relatively short-lived. Roughly a third of leaders last three years or less; half of the leaders in the data endure less than six years; and two-thirds of leaders do not last ten years. The large number of short time-series for the dynamic panels means that for relatively short-lived leaders, the dynamic and the static models would likely yield very similar estimates since there is little “time” for the priors drawn from a distribution pegged at the lagged value of the dependent variable to depart from the initial prior for each leader, which is drawn from a distribution pegged at 0.<sup>6</sup> If one were interested estimating a dynamic model with this data, we would encourage mixing (short) panels treated as locally independent with (longer) panels modeled dynamically to test whether this mixed model yield statistically different estimates than a static IRT model, as implemented for the latent estimate of personalism discussed here.

Despite the possibility that the independence may be violated for some indicators, we emphasize that the applications using the IRT estimate can address serial dependence by explicitly model non-independence within leaders or regimes using clustered errors.<sup>7</sup> In data applications that use the estimated latent trait, we can directly model panel heterogeneity among distinct leaders (or regimes) to isolate variation over time within them.<sup>8</sup> Such an applied modeling choice attempts to deal with non-independent data in applications that use personalism as a dependent or independent variable.

### 3.1.3 IRT estimates

In this section we discuss IRT estimates of a latent trait we believe corresponds to the concept of personalism in autocracies. We examine estimates from three model specifications: using the 11 items listed in Table C-1; using 10 items, while omitting the item with the lowest discrimination parameter estimate, `heirclan`); and using 8 items, while dropping `heirclan`, `heirfamily`, and `leaderrelative`.

Figure C-2 shows the item characteristic curves (ICC) for the items used in the 11-item model. The ICCs maps how well items distinguish among observations along the values of the latent trait. A steep curve indicates that the item does well at splitting high and low *personalism* cases (assuming an accurate model). This means the item provides substantial information about which cases are highly personalist and which ones are not. A shallow curve indicates that the item does not provide substantial information to distinguish cases along the latent trait. Figure C-2 highlights the three items with the smallest slopes: `heirclan`, `PartyExcomPers`, and `partyrbrstamp`. The

---

<sup>6</sup>Unsurprisingly, leaders that last longer (i.e. more than five years) have higher personalism scores (at least as measured assuming local independence) than do short-lived leaders. The 244 leaders that last less than five years have an average personalism score of 0.19 while the 261 leaders that last longer have an average personalism score of 0.40. This difference is statistically significant at the 0.001 level (two-tailed test with 505 leader-observations).

<sup>7</sup>The IRT model can incorporate dependence by clustering on leader in the estimates of the latent trait.

<sup>8</sup>One drawback to using a dynamic model in this application is that this approach, by design, smooths the estimated latent variable over time within units. This means, for example, that an increase in personalism in year  $t$  based on the historical evidence will increase personalism in year  $t-1$ . The smoothed estimates will therefore be less precise in picking up sharp changes in personalism in the leader time-series.

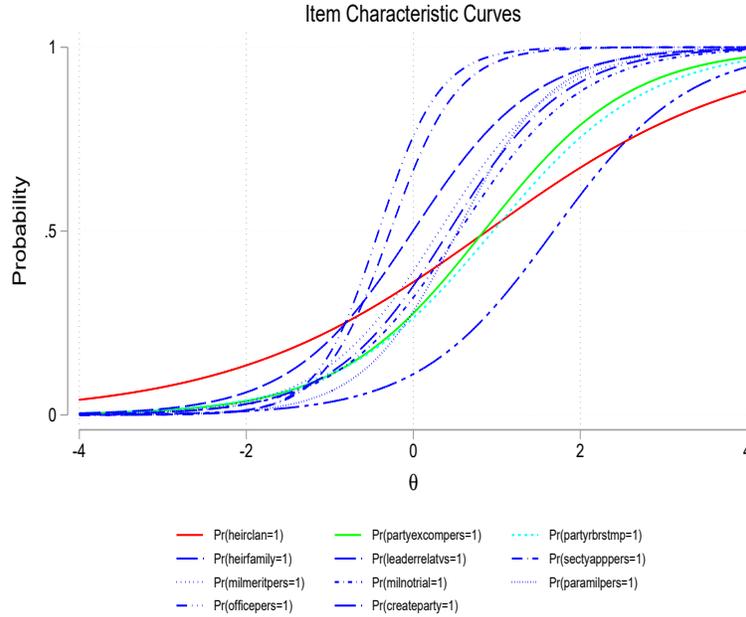


Figure C-2: *Item characteristic curves, 11-item model.*

item `heirclan` has the smallest estimate of the slope, indicating it provides the least information in distinguishing cases. This evidence corresponds to the (relatively) weak loading of `heirclan` on the first dimension shown in Figure C-1. As noted above, `heirclan` also provides substantial information on a possible second dimension. For these reasons, we re-estimate the IRT model, while omitting this item.

Figure C-3 shows the item information functions from the 10-item IRT. The height of the curves, shown on the vertical axis (Information), indicates how well the item separates high and low personalism cases. The items `officerspers` and `SectyappPers` have the highest curves (i.e. largest estimates of the “discrimination” parameter and thus do the best at splitting cases into different values of the latent trait).

The placement of the curve along the horizontal axis,  $\theta$ , shows the location along the values of the estimated latent trait where the item places the high and low cases. This measure of “difficulty” can be interpreted such that difficult items, with curves placed further to the right on the horizontal axis, are those for which an observation must be highly personalist to observe a 1 for this item. The items `officerspers` and `SectyappPers` are less “difficult” insofar as a case only has to be slightly personalist for the observations to score a 1 on these items. In the data, these items are observed frequently, as indicated by their large mean values shown in Table C-1. Similarly, `createparty` is observed least frequently, and the IRT model estimates its “difficulty” parameter as the largest of all the items (i.e. furthest to the right on the horizontal axis).

The plots in Figure C-3 indicate that the information contained in the items splits the observations all along the latent space, as can be observed by noting that the peaks of the IIFs are spread across different values of  $\theta$ . This means the items have different “difficulty” parameter estimates across most of the latent space.

Next we iteratively drop the two items that have the lowest estimates of the discrimination

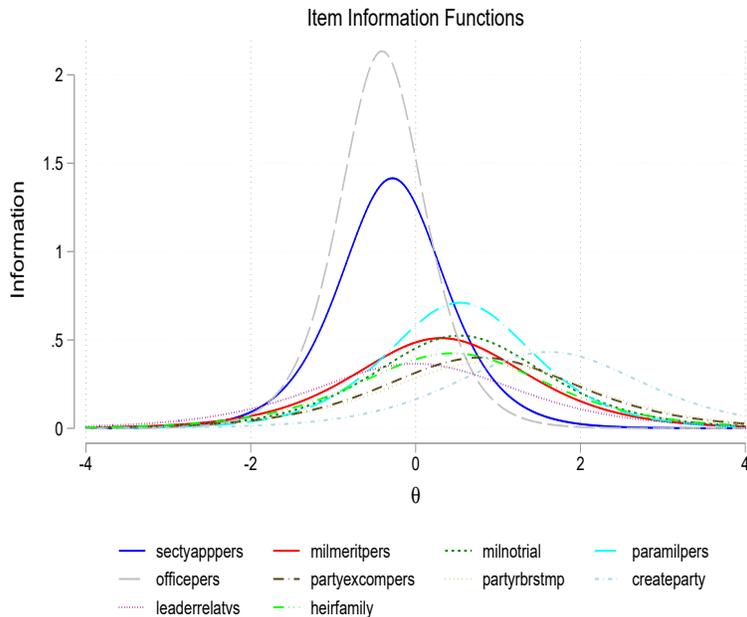


Figure C-3: *Item information functions, 10-item model.*

parameter ( $\beta$ ): `heirfamily` and `leaderrelative`. Based on visual analysis of the Figure C-1, these two items are associated with a possible second dimension in the data that we believe, based on knowledge of the how these items were coded, captures informal practices of politics in autocracies rather than having a support party that structures these interactions.<sup>9</sup>

Figure P-4 plots the item information functions (IIF) for the eight items that contribute to the latent estimate of *personalism*, or  $\theta$ .<sup>10</sup> The vertical axis measures the item discrimination parameter: higher values indicate more information in the latent estimate over a smaller range of  $\theta$  values. The horizontal axis corresponds to the “difficulty” parameter: larger values indicate items for which observations have a higher estimate of  $\theta$ . If the IRT model is accurately estimating latent *personalism*, more “difficult” items are those for which an observation must be highly *personalist* to observe a 1 for this item. This parameter captures how well an item splits high and low *personalism* cases at a particular point in the latent space.

The item `officepers` is the most discriminating, while `paramilpers` is the least. The items `officepers` and `sectyappers` are the least “difficult” (i.e. placed furthest to the left on the horizontal axis) while `createparty` is the most “difficult”. Again, the plot shows that the items split observations all along the latent space, as can be observed by noting that the peaks of the IIFs are spread across different values of  $\theta$ , and not just bunched up around a fixed value such as

<sup>9</sup>We can also use information criteria to select model specification. Comparing BICs after dropping items iteratively also shows that `heirfamily` and `leaderrelative` contribute the least to model fit.

<sup>10</sup>While not shown here, the test characteristics curve indicates that 95 percent of the observations fall between 0 and 7, meaning very few observations are positive for all eight items that contribute to the latent estimate.

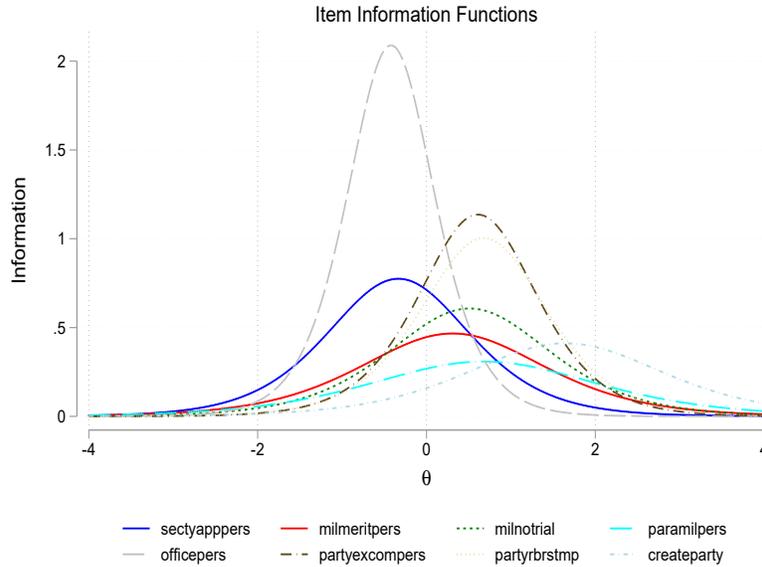


Figure P-4: *Item information functions, 8-item model.*

0.

### 3.2 Assessing the personalism index

In this section, we compare the latent estimate of personalism to existing efforts to measure the same concept: the categorical indicator of personalist regime introduced by Geddes (1999; 2003) and updated in Geddes, Wright and Frantz (2014) and Weeks (2012). The latter builds on the former by using the original coding details that Geddes (1999) employed to measure categorical personalism and updating information for some specific leaders. We include four estimates of personalism from the time-varying data introduced in this paper: the estimate produced by the EFA model in the main text (**Factor**) and the three IRT-2PL estimates discussed in this appendix.

Figure C-5 displays the pairwise correlations for the different measures.<sup>11</sup> The first column shows the correlations between the categorical GWF variable for personalist regime. It is correlated with the Weeks measure at 0.56 and the various latent estimates at roughly 0.40. The second column shows the correlations between the Weeks measure and the other measures; it is correlated with the latent measures at between 0.57 and 0.63. The latent estimate from the factor model is most strongly correlated with the 11-item IRT-2PL estimate (0.90) and less strongly correlated with the 8-item IRT-2PL estimate (0.78). The 10-item and 11-item 2PL estimates are very highly correlated, while the 8-item estimate is correlated at 0.92 and 0.94 with the 10-item and 11-item estimates.

The correlations therefore suggest that various methods of combining the items into a latent index (IRT with 8 items, IRT with 11 item, factor analysis) yield similar estimates of the level of personalism. Further, the latent measures are not particularly highly correlated with extant measures, such as the categorical GWF variable and the Weeks data.

<sup>11</sup>Correlations between various GWF variables are for 4,591 observations; correlations with Weeks are based on 2,553 observations.

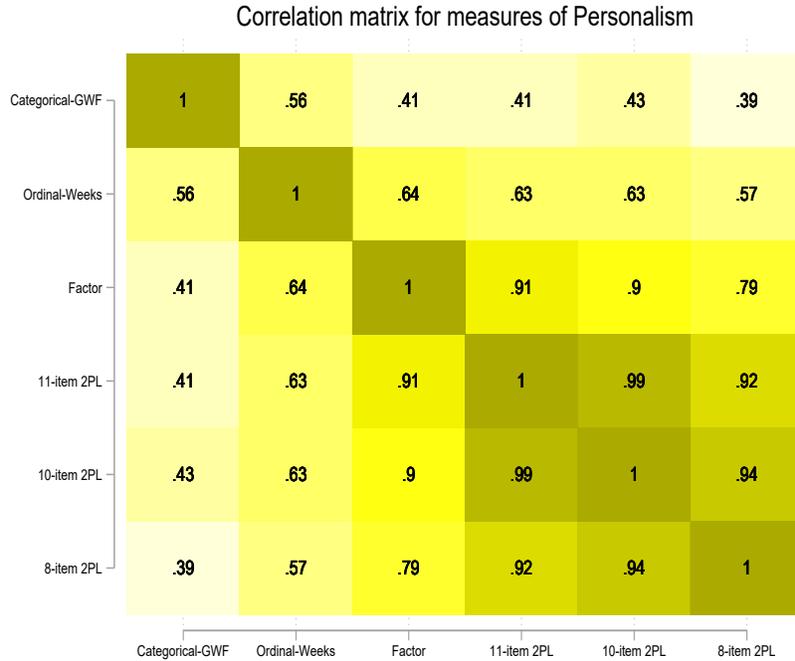


Figure C-5: *Correlation among personalism measures.*

Next, in Figure C-7, we compare the “within” variance for measures of personalism: the three IRT latent measures (IRT), the third dimension from the factor model (EFA) and the Weeks measure. We calculate variance based on regime-cases and individual leaders as the cross-section units. Note that the categorical GWF measure of personalist regime is *time-invariant*; so bear in mind that the “within” variance is 0 for this variable – on both axes. The vertical axis depicts the ratio of “within” to “total” variance<sup>12</sup> for different cross-section units (regime and leader), while the horizontal axis corresponds to the “within” variance. Higher values on the vertical axis indicate that more of the total variation in the variable is contained within the cross-section unit (e.g. over time within a regime or leader) rather than between cross-section units. This comparison is important because in many applications we will want to use the personalism measure in a statistical model that isolates variation over time within a cross-section unit (i.e. fixed or random effects models).

The left plot in Figure C-7 shows that between 30 and 40 percent of the variation in the Latent measures (IRTs and EFA) is within cases, as depicted on the vertical axis (and 60-65 percent is between cases). In contrast, 19 percent of the variation in with Weeks measures is within-case (and none of the variance in the GWF categorical indicator is within-case). Further, the latent estimates have more “within” variation overall (i.e. higher values on the horizontal axis) and not just as a share of total variation. This pattern is repeated in the right plot, which decomposes variation by leader. For example, the IRT-8 measure has within-leader variance of 0.355 and the Weeks measure has within-leader of 0.045. Finally, these plots show that the latent estimate from the 8-item IRT

<sup>12</sup>Letting  $N$  be the total number of observations,  $k$  be the number of groups,  $n$  be the number of observations within a group,  $\bar{x}$  be the group mean, and  $\bar{\bar{x}}$  be the grand mean, then the “within” variance is:  $\frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{N-k}$ . Total variance, using this notation, is  $\frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2}{N-1}$ .

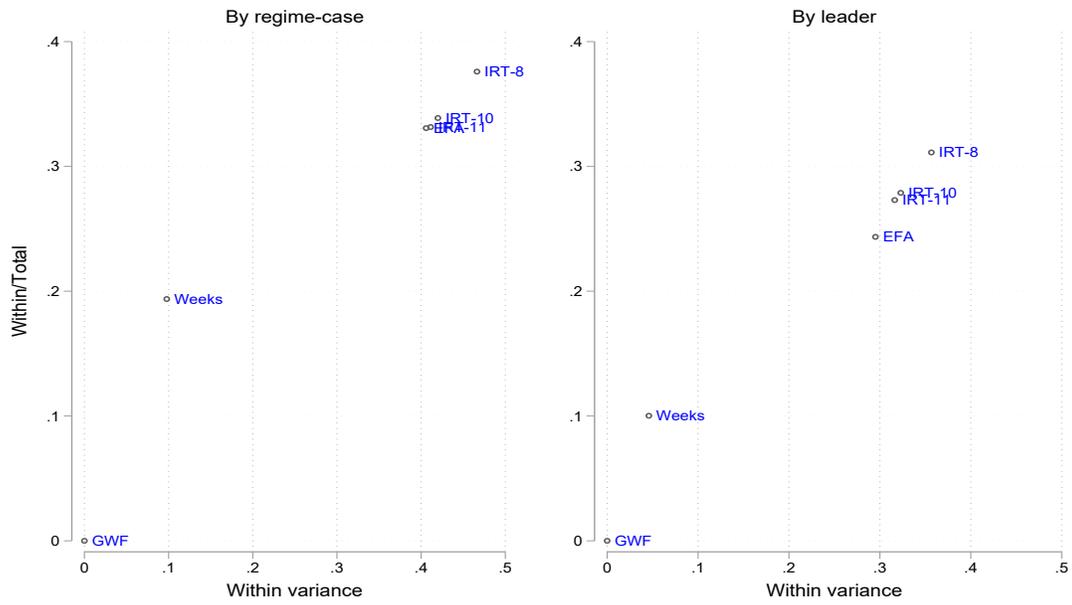


Figure C-6: *Variance decomposition.*

2PL model has more “within” variation than the other latent estimates (EFA, 11-item IRT-2PL, and 10-item IRT-2PL). Overall this plot suggests that the presence of distinct information about the level of personalism, not captured in prior measures of this concept.

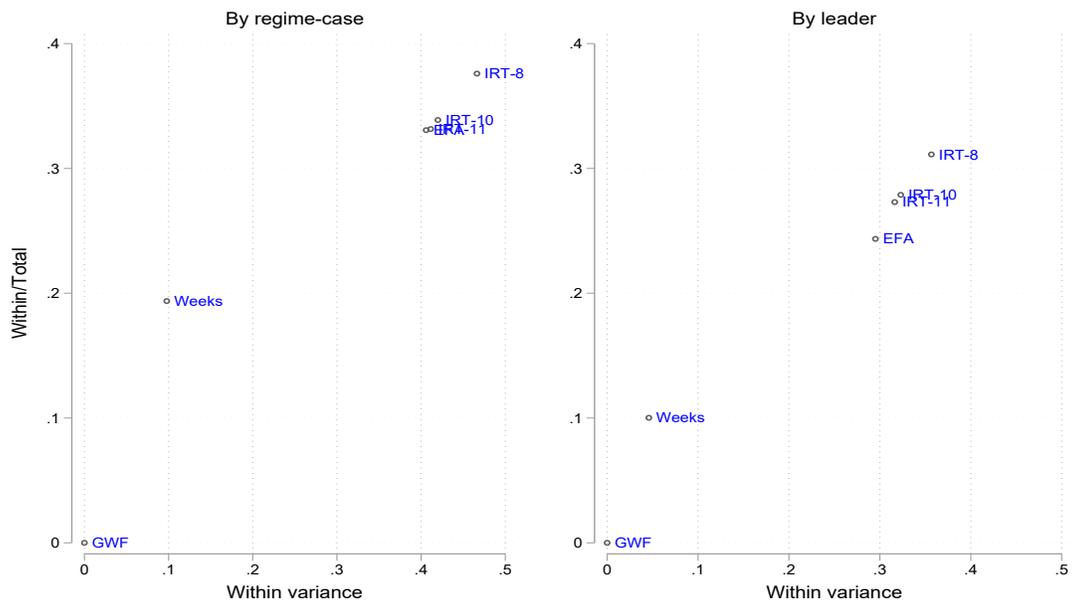


Figure C-7: *Variance decomposition.*

As a final illustration of the difference among these measures in their respective “within” variation, we return to the Chinese Communist Party (CCP) example discussed in the main text, in particular the rise and fall of Mao’s power. Figure C-8 shows various personalism measures for this regime. Recall that the GWF categorical measure of personalist regime codes the entire period, including Mao’s rule, as a ‘dominant party’ regime; thus the personalism score would be 0 for all years of CCP rule. First we note that the Weeks measure improves upon the GWF categorical indicator by distinguishing Mao as more personalist than his successors. The latent measures (Factor and the 8-item IRT-2PL) capture the rise and fall of Mao’s personal power. They also document Deng’s behind the scenes power in the early 1990s. Finally, we can see that the estimate from the EFA, shown as *Factor* in the plot, has less variation over time than the 8-item IRT estimate.

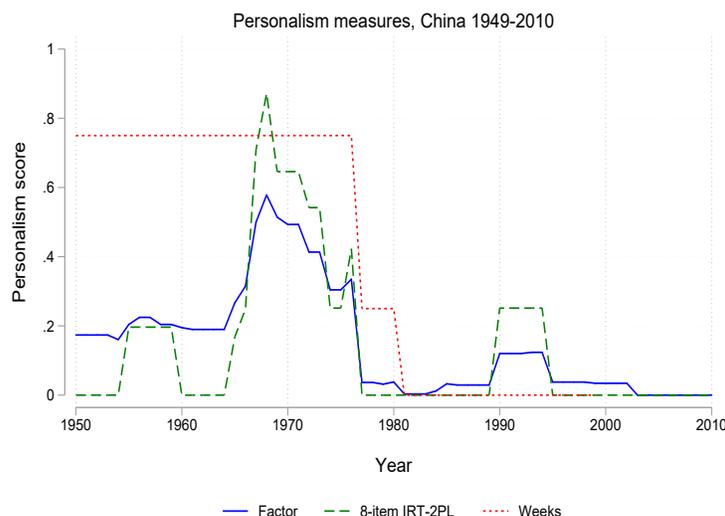


Figure C-8: *Measures of personalism for China.*

### 3.2.1 Comparison with Varieties of Democracy data

This section compares the personalism index (*IRT-Pers*) – as well as the party and military dimensions identified in the exploratory factor analysis in the main text – with data from the Varieties of Democracy (V-Dem) project on parties, the military, and executive constraint. Figure C-9 show the correlation matrix for these measures and the V-Dem measures of party institutionalization and investigations. *IRT-Pers* is negatively correlated with all the V-Dem measures, with the largest (negative) correlation with executive respects the constitution. This is consistent with the prior expectation that personalization “hollows out” the power of formal institutions, such as the constitution, in autocracies. That said, executive respects the constitution is also negatively correlated with the Military dimension, which suggests that respect for the constitution measures may simply be picking up leader coup entry. The Party dimension is positively correlated with V-Dem measures of party institutionalization, as expected.

Figure C-10 shows the correlation with V-Dem measures of how the executive gained power (path), how he is removed, and whether he needs military, party, or no approval. We should expect personalism to have little relationship with these measures, except perhaps “executive no approval”.

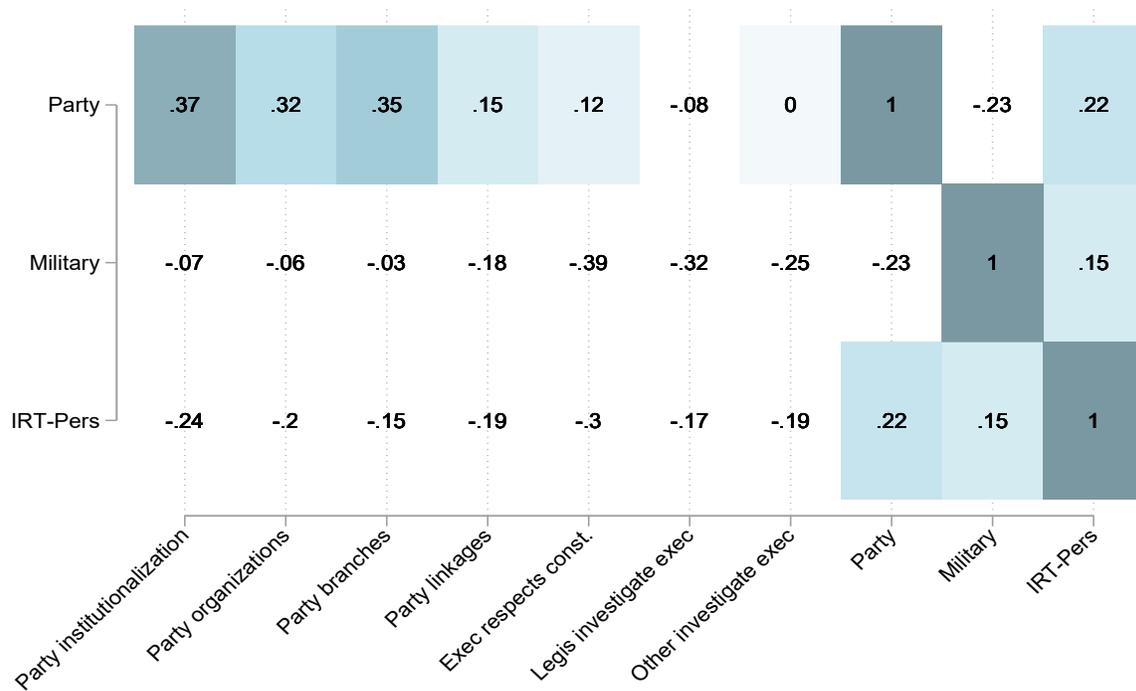


Figure C-9: *Comparison with V-Dem measures of party institutionalization and investigations.*

Indeed, *IRT-Pers* is only positively correlated with “military removal” and, more strongly, with “executive no approval”. The Party dimension is relatively highly correlated with party removal from power and party approval, while the Military dimension is relatively highly correlated with military approval. These latter findings speak to convergent validity for the Military and Party dimensions.

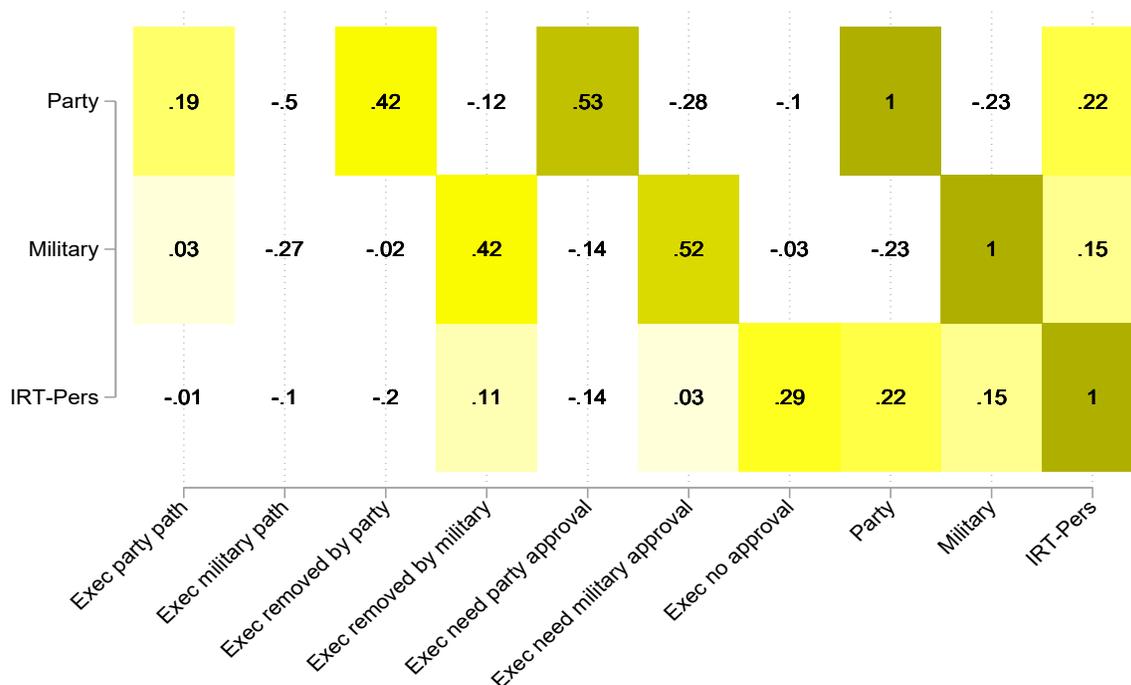


Figure C-10: *Comparison with V-Dem measures of removal and approval.*

Overall, the V-Dem measure most closely corresponding to personalism is “executive no approval”, which provides some evidence of convergent validity for this measure. However, this correlation is only 0.29, indicating that most of the variation in *IRT-Pers* is not picked up in this V-Dem variable.

### 3.2.2 Reliability

This subsection discusses the reliability of the personalist index. One form of reliability is test-retest reliability, which entails testing the same group at different times. Conceptually, this approach is nonsensical in this application because the measure is constructed to pick up changes over time in regimes, not to measure the regime at various points in time in an effort get an “average” measure of personalism for each regime. Indeed the latter is a personalist regime dummy variable.

That said, we can think differently about testing the units at different times by estimating the IRT model using data from different *calendar time* periods, calculating the estimated latent variable using all the data but different model parameters, and comparing the resulting latent variables. For example, we first divide the sample at the median calendar year, 1980, into two groups, pre- and post-1980, and then estimate the model parameters using data from each bin separately. We

then calculate the latent measure (using all data) with each set of parameters and finally compare the variable computing from each model. This approach tells us whether manifest items contribute similar (or distinct) types and size of information to the variable estimate in different time periods. There could be bias, for example, stemming from researchers having more difficulty tracking down information on one (or more) item(s) in earlier rather than later periods. This would make the measure less reliable for earlier periods.

The first test in Figure C-11 shows the results from the procedure. The data from the Group 1 model – derived from a model using only pre-1980 data – are correlated with the latent measure at greater than 0.99; the same is true for the Group 2 data, which are derived from a model using only post-1980 data. The second test divides the data points (4,591 regime-year observations) into two random bins and conducts the same test: again the correlations are incredibly high.

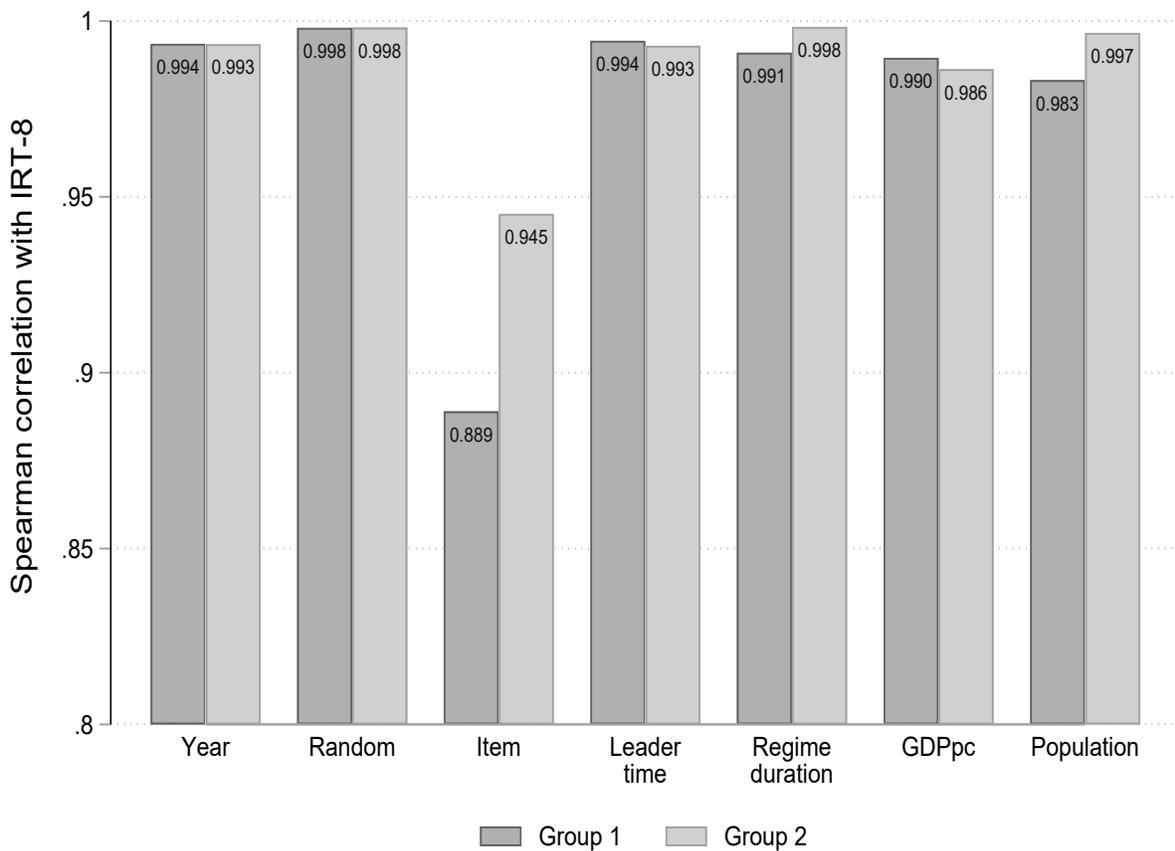


Figure C-11: *Reliability tests.*

Next, we separate the eight manifest *items* into two bins; because there may be two sub-components to personalism (party and military) we do not do this randomly but rather ensure that each of the two bin contains two items related to the military and two items related to the party. We then test an IRT model for each group or bin of items and calculate the predicted latent personalism index from each model. The third set of bars in Figure C-11 show this result. The

measure estimated from one group of items is correlated with personalism measure at 0.89 while the measure from the other group is correlated with the personalism measure at almost 0.95.

The last four reliability tests divide the sample at the median of an important variable to test whether the model output is similar when constructed from data using observations below the median value on a variable and when constructed from a second group of data points including observations with values above the median. For example, the fourth test in Figure C-11 shows the correlation between the latent personalism index and measures derived from data above (Group 1) and below (Group 2) the median leader time in power (which is seven years). Both correlations are above 0.99. We repeat this exercise three more times for three different variables: regime duration, GDP per capita, and Population size. In all these tests, the correlations are above 0.98. There could be more information in the written record about informal politics and personalism in China, for example, than in Gabon, which might produce unreliable estimates for personalism in small countries such as Gabon. Because the model estimates are relatively stable, yielding roughly the same measures of personalism when derived from different samples, we have less concern about this type of unreliability.

Overall, these tests indicate a highly reliable measure. The parameters of the IRT model – which when combined with the observed data yield the estimates of the latent personalism index – are stable.

### 3.2.3 Validity

Conceptually, personalism in this project is: the leader’s concentrated power at the expense of his closest supporters organized around two collective action organs: the military and the party (if there is one). **Content validity** is assessed in the discussion of Appendix Figure C-1 (above), which shows how the items used in the latent personalism measure might be further distinguished along two sub-components of personalism relating to the two key collective action organizations most dictatorial leaders face: the military/security apparatus and the supporting political party.

One way to assess **face validity** is to plot the change over time within countries in the personalism index to trace the historical events that are observed and recorded for use as manifest indicators of personalism. We have done this for China in the main text. We refer readers to ? for a process tracing account of the political events in North Korea that are measured in the personalism index.

Figure C-5 speaks to (concurrent) **convergent validity** by showing that the Personalism index is correlated with GWF and Weeks measures of personalist regime, which are conceptually similar but do not vary substantially within leader or regime.

Table C-2: Spearman correlations with latent personalism index

	<u>overall</u>	<u>within</u>	<u>between</u>
Personalist regime (GWF)	.388	.	.462
Personalist rating (Weeks)	.569	.213	.647
Personalist military (Svolik)	.198	.185	.17
Corporate military (Svolik)	.04	-.039	.053
Executive constraint (Polity)	-.301	-.032	-.244

Table C-2 shows the overall, within, and between correlations for these measures of personalism

and related concepts. In this table, the *overall* Spearman correlation reflects both within- and between-unit variation, where the cross-section unit is the *regime*. The latent personalist index is positively correlated overall with extant – but less fine-grained – variables measuring the same concept. Further, while the correlation is low with the military variables, the correlations between the personalism index and the personalist military are higher than those between the index and corporate military, which provides some (relative) evidence of **convergent validity**. The personalism index is also negatively correlated with a measure of executive constraint, as expected, helping to establish **discriminant validity**. Finally, the between-unit correlations are generally higher than the within correlations, which demonstrates **discriminant validity**: the data for the personalism index was coded (and structured) with the intent to capture more time-varying information than exists in extant data.

**Discriminant validity** is further established by showing low correlation between the third (personalist) dimension produced by the exploratory factor analysis and variables measures aspects of democracy from the Polity and V-Dem projects (Table B-2, Figures B-3 to B-7). The third (personalist) dimension produced by the exploratory factor analysis is also *not* highly correlated with extant measures of political institutions and the military (Gandhi, Hadenius and Teorrel, Svolik, and DPI), as shown in table B-3. Table C-3 continues in this spirit by showing the within-unit correlation with these extant variables measuring concepts related to political institutions and the military in dictatorships. All of these correlations are less than 0.20 (in absolute value), indicating **discriminant validity**.

Table C-3: Within correlations with latent personalism index

Parties (Gandhi)	0
Parties (H&T)	-.054
Parties (DPI)	.132
Parties (Svolik)	.002
Party (Gandhi)	.112
Party (H&T)	.184
Party (DPI)	.176
Party (Svolik)	.117
Institutions (Gandhi)	.077
Legislative index (Svolik)	.082
Legislative index (DPI)	.139
Military (Gandhi)	.136
Military (H&T)	-.13
Military (DPI)	.095
Military (Svolik)	.153
Corporate military	-.039
Personalist military	.185
Indirect military	-.023

Final look at discriminant validity is to compare the personalism index with data on potentially related concepts from the V-Dem project. Figures C-9 and C-10 compare the measure of personalism to measures of party institutionalization, investigations into the executive, executive removal, and executive approval. None of these measures from the V-Dem project correlate highly with the

personalism measure. The measures that have the strongest (in absolute value) correlation with personalism are “executive requires no approval [from the military or party]” (0.29) and “executive respects the constitution” (-.30). These are two concepts most closely related to personalism, so this is not surprising. Nonetheless, these correlations remain relatively low and thus indicate that the personalism data is relative unique.

### **3.3 Conclusion**

In sum, this Appendix presents and discusses an IRT approach to estimating the latent level of personalism in dictatorships. Should readers want to use the data presented in the paper for their own applications, we recommend the 8-item IRT 2PL estimate. However, no measure is perfect and we encourage interested readers to employ the data to estimate a measure of personalism best suited for their research application.

## 4 Appendix D: Additional tests for MID initiation

### 4.1 Details of MID initiation results in main text

This Appendix provides further details on the tests for MID initiation presented in the main text. The main text provides results from three tests for ordinary logit and FE logit estimators. The results are reported in Table D-1, with standard errors in parantheses. The first (logit) and fourth (FE) columns are exact reproductions of the tests reported in Weeks (2014, Table 2.3). Columns 2 and 5 report results from tests that slightly alter the sample because there are a handful of country-year observations that are coded by Weeks (2014) and Geddes (2003) as autocratic but are coded as non-autocratic (on January 1) in the new data: Czechoslovakia 1990; Ecuador 1967; Haiti 1957; Honduras 1972; Thailand 1989; and Uruguay 1985. Omitting these observations reduces the sample size by 0.3 percent but, as shown in the D-1, this does not alter the result. Finally, the third and sixth columns report results using the reduced sample by substituting *G-pers* for *W-pers* (and  $G-pers \times W-mil$  for  $W-pers \times W-mil$ ).

Figure D-1 plots the marginal effects of personalism, as estimated from the FE models in columns 4 (original) and 6 (new data). Substantively, the original FE model indicates that personalism increases the risk of MID initiation by between 23 percent (0 on the militarism index) and 13 percent (1 on the militarism index); in contrast the same model with *G-pers* suggests personalism increases MID initiation by between 9 percent and 8 percent. The estimated marginal effect of the latter is less than half the original estimated effect.

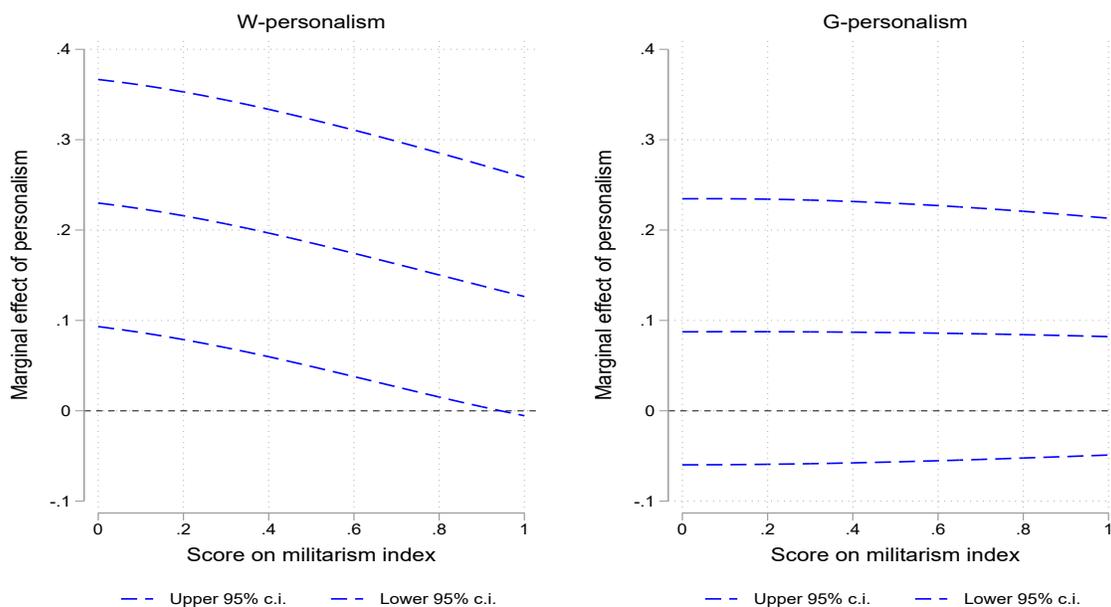


Figure D-1: *Marginal effects of personalism on MID initiation, two measures.*

Table D-1: MID Initiation

	(1)	(2)	(3)	(4)	(5)	(6)
W-pers	1.076*	1.079*		1.116*	1.150*	
	(0.23)	(0.23)		(0.33)	(0.33)	
G-pers			1.392*			0.389
			(0.25)			(0.34)
W-pers $\times$ W-mil	-0.809*	-0.806*		-1.293*	-1.360*	
	(0.31)	(0.31)		(0.61)	(0.62)	
W-pers $\times$ W-mil			-0.977*			-0.103
			(0.42)			(0.65)
W-mil	1.256*	1.253*	1.194*	1.484*	1.566*	0.654
	(0.28)	(0.28)	(0.27)	(0.53)	(0.53)	(0.47)
democracy_2_lag	0.475*	0.474*	0.504*	-0.156	-0.155	-0.165
	(0.15)	(0.15)	(0.15)	(0.24)	(0.24)	(0.24)
cap_1_lag	12.723*	12.698*	7.639	1.553	1.740	-5.353
	(5.35)	(5.36)	(5.34)	(10.68)	(10.69)	(10.25)
cap_2_lag	3.739*	3.635*	3.825*	9.170*	8.958*	9.375*
	(1.77)	(1.79)	(1.75)	(2.91)	(2.92)	(2.88)
initshare_lag	-0.037	-0.038	-0.115	0.263	0.258	0.374
	(0.21)	(0.21)	(0.21)	(1.05)	(1.05)	(1.06)
dependlow_lag	16.057*	16.081*	13.883*	27.719	27.751	26.678
	(3.58)	(3.60)	(4.24)	(17.16)	(17.22)	(16.33)
majmaj_lag	0.613	0.619	1.413			
	(0.87)	(0.87)	(0.89)			
minmaj_lag	1.341*	1.340*	1.308*			
	(0.28)	(0.28)	(0.27)			
majmin_lag	-0.230	-0.233	0.600			
	(0.70)	(0.70)	(0.69)			
contigdum_lag	2.232*	2.242*	2.267*			
	(0.62)	(0.62)	(0.60)			
logdist_lag	-0.157*	-0.156*	-0.152*			
	(0.07)	(0.07)	(0.07)			
s_wt_glo_lag	-0.040	-0.055	-0.000	0.002	0.037	-0.167
	(0.31)	(0.31)	(0.30)	(0.51)	(0.51)	(0.50)
s_lead_1_lag	-0.491	-0.493	-0.462	-1.742*	-1.767*	-1.512*
	(0.37)	(0.37)	(0.36)	(0.58)	(0.58)	(0.57)
s_lead_2_lag	0.949*	0.939*	0.941*	0.347	0.392	0.033
	(0.26)	(0.26)	(0.26)	(0.57)	(0.57)	(0.56)
(Intercept)	-4.835*	-4.831*	-4.881*			
	(0.72)	(0.72)	(0.71)			
N $\times$ T	268,458	267,656	267,656	8178	8158	8158
Directed dyads	12885	12879	12879	265	264	264

## 4.2 Robustness tests

### 4.2.1 Compare autocracies with democracies

Next, we present four sets of tests that further examine the findings for personalism. First, we conduct similar sets of tests to those reported in Figure 8 of the main text and Table D-1 but extend the sample to allow comparison of autocracies with democracies. Figure D-2 reports the results. Estimates shown with the diamond ( $\diamond$ ) in the left panel are an exact reproduction of the ordinary logit results in Table 2.3 column 3 in Weeks (2014), while estimates shown with the diamond ( $\diamond$ ) in the right panel are an exact reproduction of the FE logit results in column 4 of the original. As in the prior analysis, the estimates shown with triangles ( $\triangle$ ) use *W-pers* but with the slightly smaller sample; and estimates shown with squares ( $\square$ ) use *G-pers* with the reduced sample. As in the prior tests, substituting *G-pers* for *W-pers* yields almost identical results in the ordinary logit tests, but divergent results in the FE tests.

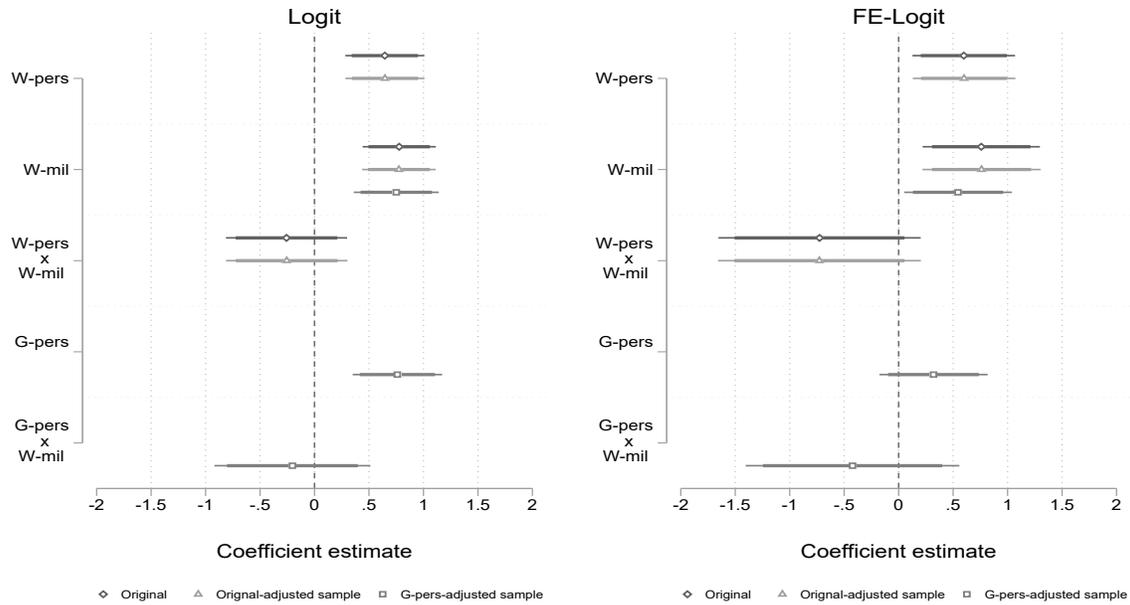


Figure D-2: *Marginal effects of personalism on MID initiation, with democracies in the sample.*

### 4.2.2 Specifications that exclude the interaction term

The second set of tests are similar to those reported in Figure 8 of the main text and Table D-1 but change the model specification to *exclude* the interaction term. This allows us to assess the average estimated marginal effect of personalism across all levels of the military index as indicated by the estimate of personalism without the interaction. Figure D-3 reports the results. Substituting *G-pers* for *W-pers* yields stronger results in the ordinary logit tests, but, again, considerably weaker results in the FE logit tests. The estimated marginal effect of *G-pers* in the FE logit model in the right panel of Figure D-3 runs from between 0.079 to 0.075, depending on the level of the militarism index. This is roughly a third to a half the size of the estimated marginal effect from the FE model using *G-pers* reported in Figure 8 of the main text and Table D-1.

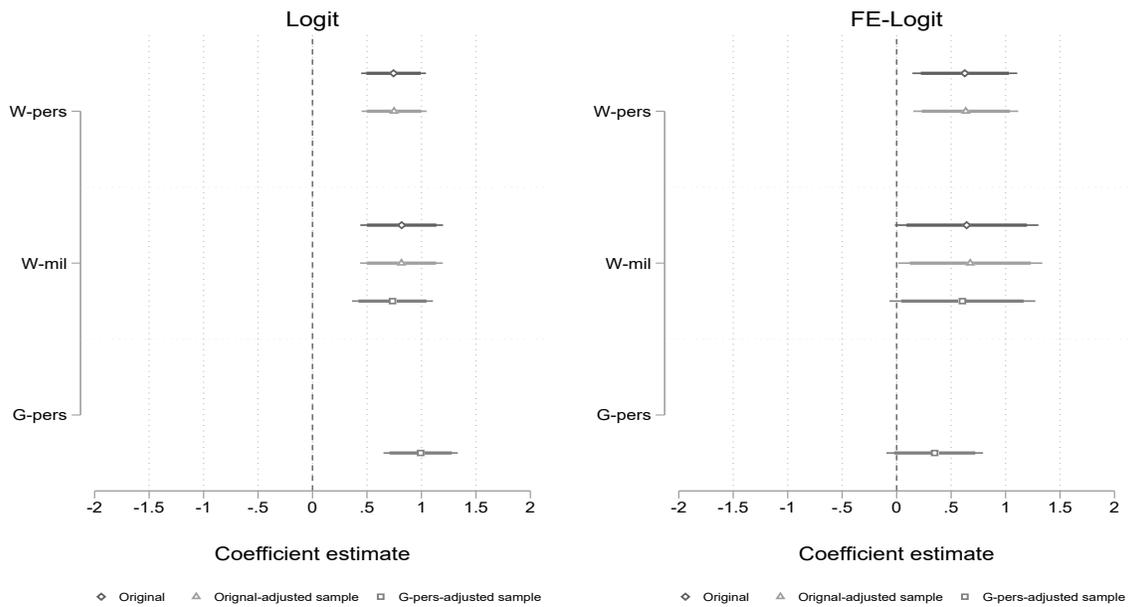


Figure D-3: *Marginal effects of personalism on MID initiation, without the interaction term.*

### 4.2.3 Testing different items in the *G-pers* measure

The third set of tests explore the FE logit result using *G-pers* but construct slightly different versions of this variable, all using a 2-parameter logistic IRT model. We test these models because we want to ensure that the divergent results for personalism in the FE logit models reported in Figure 8 of the main text and Table D-1 are not driven by using different conceptual pieces of information to construct *G-pers*.

First, we try to mimic the exact indicators Weeks employ to construct the personalism index, *W-pers*, but instead using the items from the newly collected data: `officepers`; `partyrbr`; `sectyappers`; `partyexcompers` (politburo), (heir family or clan); (military promotion or personal paramilitary); `milnotrial`; and `milconsult`. To construct a measure of personalism with these items, we use the following: `sectyappers`, `officepers`, `partyrbr`, `partyexcompers`, `paramilpers`, `milmeritpers`, `milnotrial`, `milconsult`, `heirfam`, and `heirclan`. In short, the new data have separate indicators the ‘heir apparent’ from the leader’s family and clan. We call this variable *G-pers* (*W*). Next we construct 11-item and 10-item IRT model measures. The 11-item variable, which we call *G-pers* (*11*), employs all 11 items listed in Table C-1. The 10-item variable, which we call *G-pers* (*10*), omits `heirclan`. [Recall that the *G-pers* measure only uses information from 8 indicators, as discussed in Appendix C.] Figure D-4 reports the results. None of the estimates of the personalism variable is statistically significant at the 0.10 level, though the estimate when using *G-pers* (*10*) is close. This suggests that the exact constellation of items used to construct the personalism measure does not alter the divergent findings.

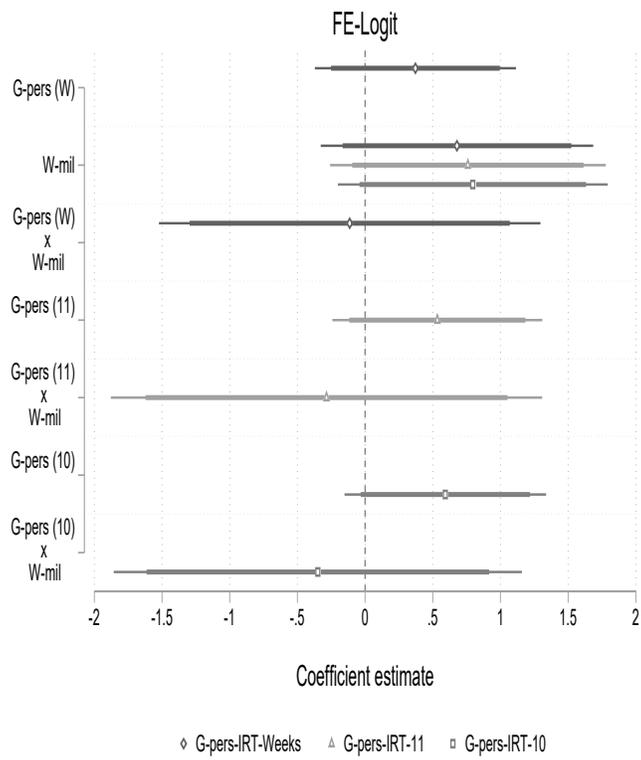


Figure D-4: *Marginal effects of personalism on MID initiation, additional G-pers measures.*

#### 4.2.4 Including early and unstable years of each regime

The final set of tests are similar to those reported in Figure 8 of the main text and Table D-1 but extend the sample to include the early years and unstable years of each regime. Geddes' original construction of the regimes data omitted the first three years of each regime from early regime classifications because so many autocracies were becoming more personalistic during these first years in power. Geddes (2003) and Weeks (2014) thus only code autocracies once they have been in power for at least three years. This omits regimes that last less than three years, and, as importantly creates holes in the data for the first three years longer-lasting regimes. The tests Weeks (2014) conduct therefore exclude these new regimes. She also excludes regimes that have previously been “unstable” according to the Polity *durable* variable. These unstable regime-years include the first three years of the Geddes' regime as well as the first three years after a major institutional change, such as the introduction of multiparty politics. Observation years when Polity *durable* variable is less than 3 are considered “unstable”.

Figure D-5 reports the results. The estimates for personalism – both *W-pers* and *G-pers* are positive and statistically significant in the ordinary logit tests, as in prior tests. Similarly, the estimate for *W-pers* is in the expected direction and size in the FE tests. But, again, the FE logit tests with *G-pers* substituted for *W-pers* yield very small estimates for *G-pers*. The marginal effect of personalism from this model ranges from 0.032 to 0.036, which are much smaller effects than those estimated in the FE logit model for *W-pers* (0.167 to 0.193).

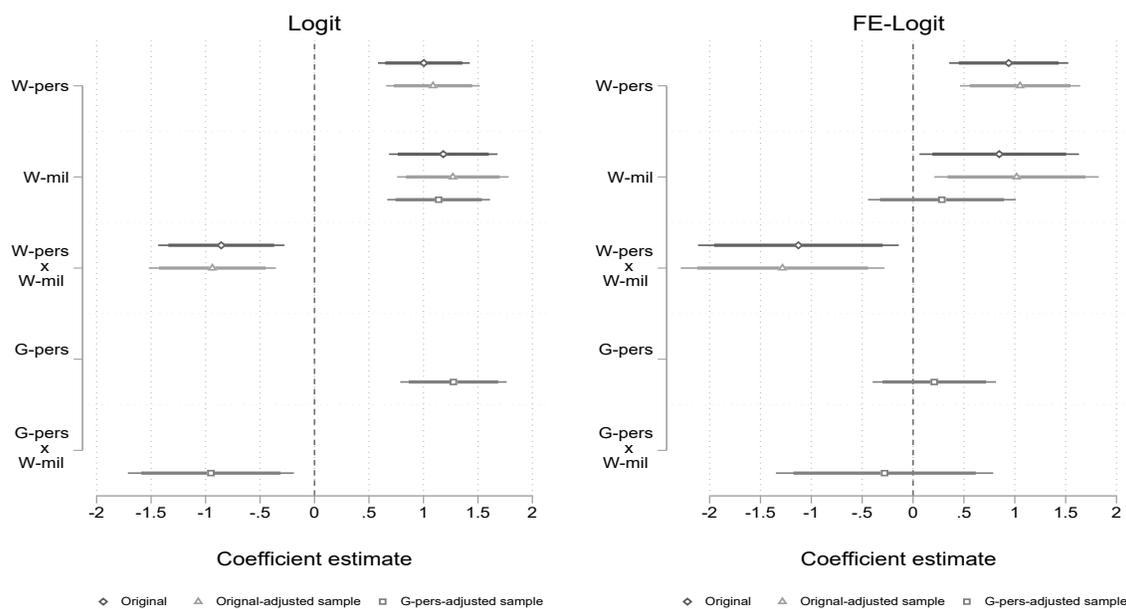


Figure D-5: *Marginal effects of personalism on MID initiation, sample extended to new/unstable regime-years.*

### 4.3 Modeling uncertainty in the latent personalism measure

Finally, we examine how the fixed effects result reported in the main text changes when we model uncertainty in the latent personalism measure. The models in the main text (Figure 9) use the point estimates of the latent variable (mean of the posterior distribution) to estimate the models but ignores the standard deviation of the posterior distribution. We generate 1000 draws for each country-year observation for personalism from a normal distribution described by the mean and standard deviation of the latent variable estimate for each country-year observation of latent personalism, derived from the IRT model. We then estimate the model 1000 times, for each draw from the distributions of the latent variables. 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 D-6 reports estimates for the main variables when accounting for uncertainty in the latent personalism measure. The estimate for  $G$ -pers is 0.069, while that for the interaction term is -0.012. These estimates are not statistically different from zero and are smaller than the estimates reported in the main text (see column six of Figure D-1).

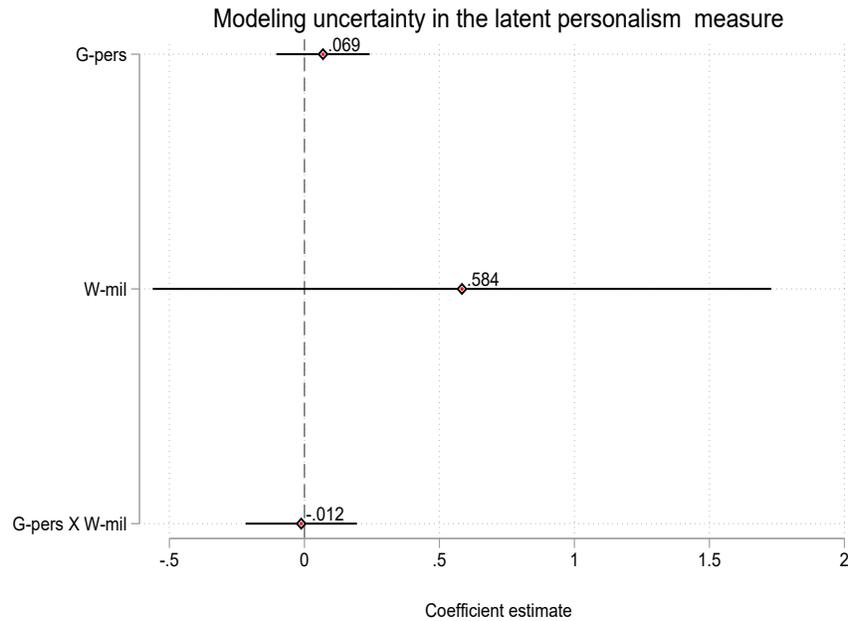


Figure D-6: *Estimates of personalism incorporating uncertainty in the latent measure of personalism.*

#### 4.4 Unpacking personalism and dispute initiation data for China and Libya

To further illustrate why the two measures produce different FE results, we next unpack the data for China and Libya. This exploration of the descriptive data illustrates both how the new personalism variable captures the rise of personalist rule over time within a leader's tenure and how this can sometimes – but not always – coincide with the relevant political outcome in an empirical application.

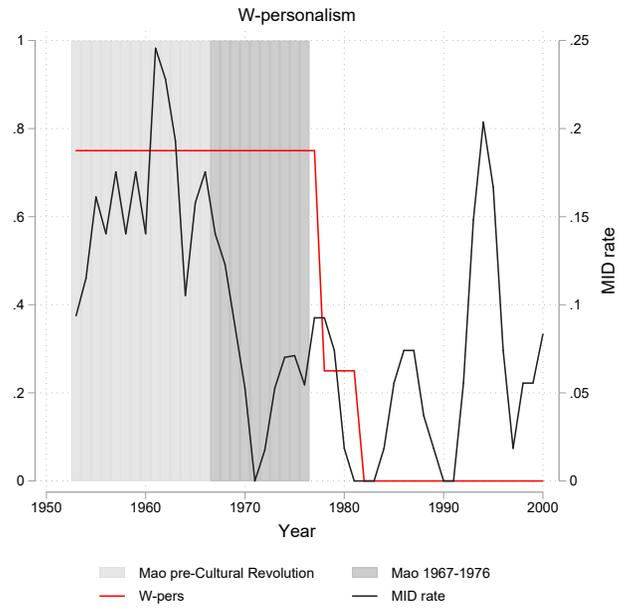
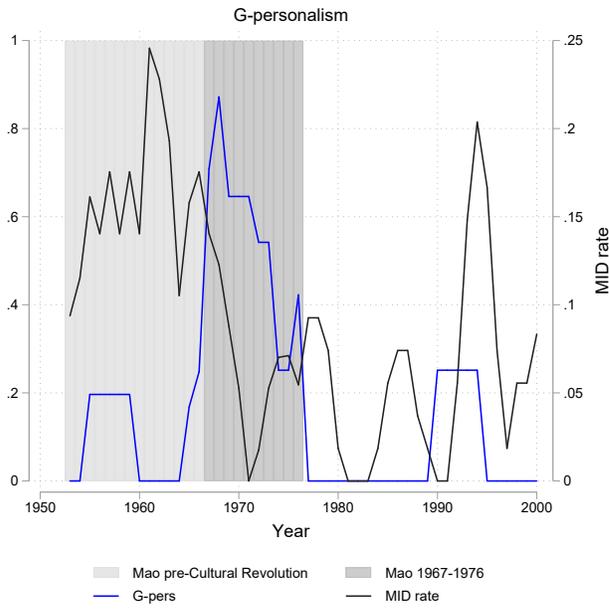
Weeks' innovation is to allow the personalism score for different leaders within the same regime to differ. For example, Mao's rule is coded as high personalism while his successors have low personalism scores. The new data on personalism improves upon this by measuring how personalism varies over time within leaders. The top left panel of Figure D-7 shows the trajectory of personalism in China, as measured by *G-pers*, with the blue line: personalism is relatively low from the 1950s into the mid-1960s and then rises with the start of the Cultural Revolution in 1966; and post-Mao leaders have low personalism scores. The plot also shows the MID-initiation rate (per year) for China with the black line, while the shaded bars distinguish the early period of Mao's rule from the later period of his rule when the personalism score is higher.

The baseline MID initiation rate for China prior to 1966 is 16 percent; this number drops to just over six percent during the 1967-1976 period and to just under six percent in the post-Mao period. These baseline MID rates indicate that Chinese dispute initiation was most common during periods of low personalism under Mao's rule. The right panel presents the same information but shows the Weeks' personalism measure, which is high during the entire period of Mao's rule and falls afterwards. High Chinese MID rates from 1953 to 1965 occur during the early period of Mao's rule, which the Weeks' data codes as highly personalist. This example illustrates the value of measuring periods of low personalism early in a leader's tenure, capturing the rise and fall of personalism over time within regimes.

The bottom panel of Figure D-7 plots similar data from Libya under Gaddafi's rule. Personalism, as measured by *G-pers*, is relatively low after the 1969 coup that brings Gaddafi to power, but rises quickly such that by the end of his first decade in power Libya has the maximum personalism score. In contrast, the Weeks' data codes this case as highly personalist throughout. In Libya, the early period of low dispute initiation coincides with a relatively low personalism score, as shown in the left plot. The period of high dispute initiation rates, from the late 1970s to the mid-1980s coincides with highly personalist rule, as shown in both plots.

These examples illustrate both how the new personalism variable captures the rise of personalist rule over time within a leader's tenure and how this can sometimes – but not always – coincide with the relevant political outcome in an empirical application. The data from China help explain why the findings from the FE models in Figure ?? diverge when substituting the new personalism measure for extant data. Further, these examples illustrate that the new time-varying measure of personalism allows researchers to test theories about how international factors – such as conflict behavior – may influence the consolidation of personalist power.

## China



## Libya

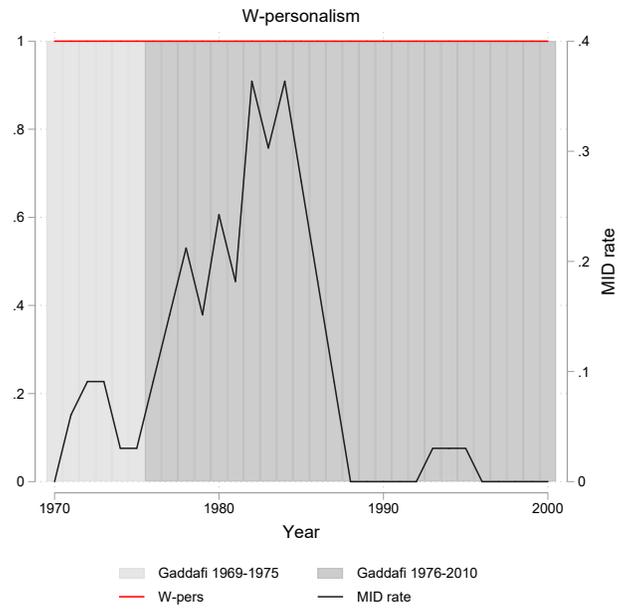
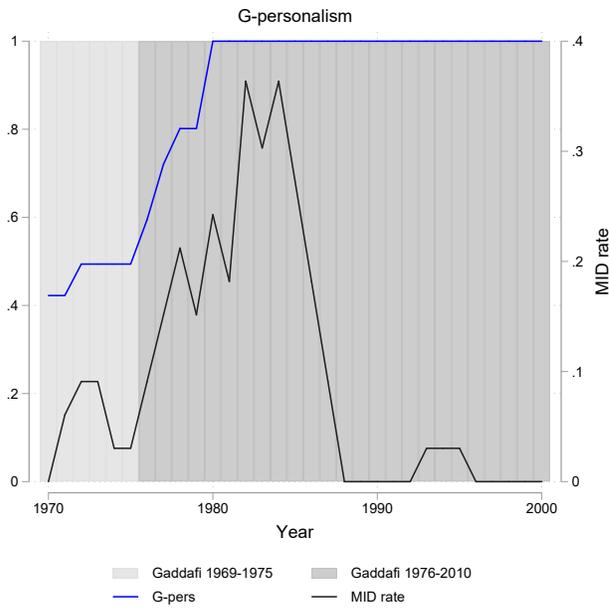


Figure D-7: Personalism and dispute initiation in China and Libya.

## 5 Appendix E: Additional tests for regime collapse analysis

### 5.1 Robustness tests

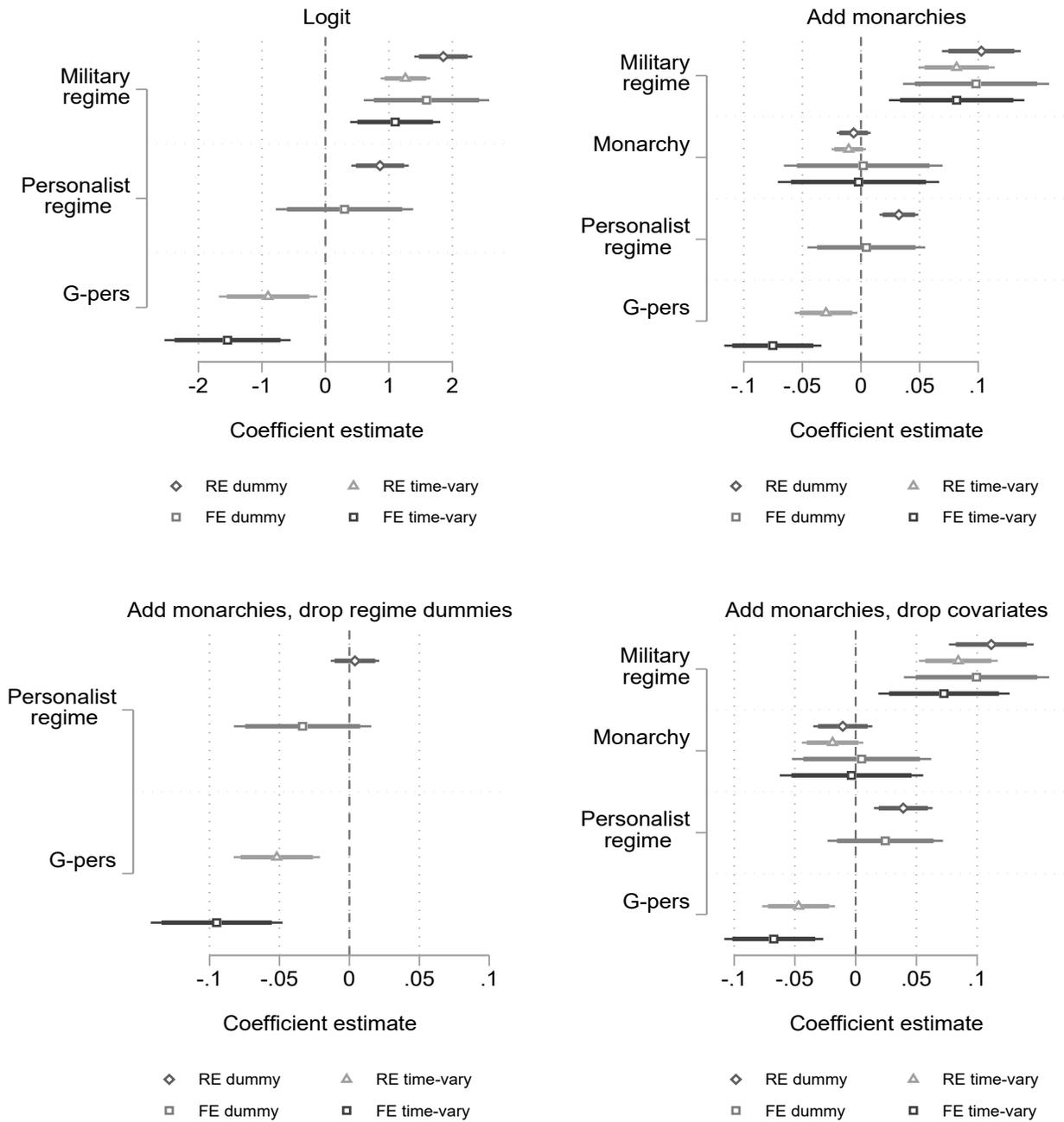


Figure E-8: *Robustness tests for regime collapse.*

## 5.2 Different types of transition outcomes

The analysis of regime collapse in the main text (Figure 11) pools all regime collapse together as the outcome. In this section, we divide regime collapses into two categories: those that result in a new democracy and those that usher in a new dictatorship. For example, the collapse of the Marcos regime in the Philippines in 1986 led to a transition to a new democracy. The Iranian Revolution of 1979, however, caused the collapse of one dictatorship (Reza Shah's regime) that led to a transition to a new, subsequent dictatorship (the theocratic regime).

In this section, we test the regime collapse model but divide these events into the two categories to examine whether personalism influences one type of collapse more than the other. Figure E-9 reports the results. A personalist regime – measured as a dummy variable – is associated with a higher likelihood of transition to new dictatorship than party regimes (blue triangles), but not significantly more likely than party regimes to transition to democracy (black diamonds). Using the continuous measure of personalism *G-pers*, however, more personalization is associated with a lower likelihood of transition to new democracy (green squares) and has little effect on the risk of transition to new dictatorship (brown squares).

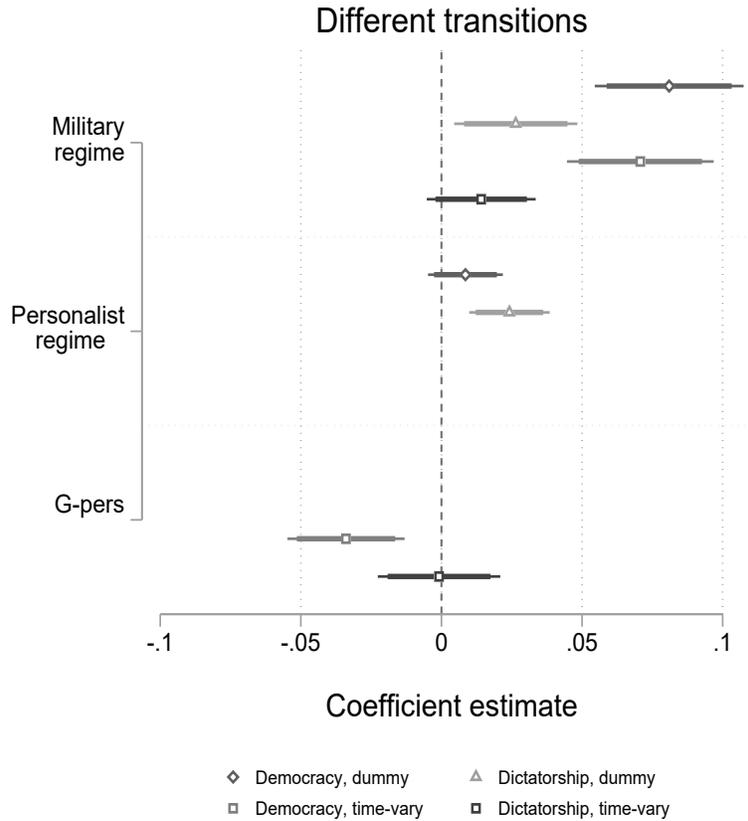


Figure E-9: *Personalism and different regime collapse transition outcomes.*

## References

- Aksoy, Deniz, David B Carter and Joseph Wright. 2012. "Terrorism in Dictatorships." *The Journal of Politics* 74(3):810–826.
- Conrad, Justin, Courtenay Conrad and Joseph K. Young. 2014. "Tyrants and Terrorism: Why some Autocrats are Terrorized while Others are Not." *International Studies Quarterly* Forthcoming.
- Davenport, Christian. 2007. "State repression and the tyrannical peace." *Journal of Peace Research* 44(4):485–504.
- Escribà-Folch, Abel. 2013. "Repression, political threats, and survival under autocracy." *International Political Science Review* 34(5):543–560.
- Fjelde, Hanne. 2010. "Generals, Dictators, and Kings Authoritarian Regimes and Civil Conflict, 1973 – 2004." *Conflict Management and Peace Science* 27(3):195–218.
- Frantz, Erica and Andrea Kendall-Taylor. 2014. "A dictator's toolkit: Understanding how co-optation affects repression in autocracies." *Journal of Peace Research* 51(3):332–346.
- Gandhi, Jennifer and Jane Lawrence Sumner. 2017. "Measuring Personalism in Non-Democracies." *International Methods Colloquium*. Slides available at: <https://bit.ly/2j0jSGT> [Accessed 10 November 2017].
- Geddes, Barbara. 1999. "What Do We Know About Democratization After Twenty Years?" *Annual Review of Political Science* 2:115–144.
- Geddes, Barbara. 2003. *Paradigms and Sand Castles*. Ann Arbor: Univ. of Michigan Press.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "New Data Set: Autocratic Breakdown and Regime Transitions." *Perspectives on Politics* 12(2):313–331.
- Gehlbach, Scott and Philip Keefer. 2012. "Private investment and the institutionalization of collective action in autocracies: ruling parties and legislatures." *Journal of Politics* 74(2):621–635.
- Hankla, Charles R and Daniel Kuthy. 2013. "Economic liberalism in illiberal regimes: Authoritarian variation and the political economy of trade." *International Studies Quarterly* 57(3):492–504.
- Kenwick, Michael. 2017. "Is Civilian Control Self-Reinforcing? A Measurement Based Analysis of Civil-Military Relations.". Working paper available at: <http://bit.ly/2xs0XJ6> [Accessed 10 November 2017].
- Leeds, Brett Ashley, Michaela Mattes and Jeremy S Vogel. 2009. "Interests, institutions, and the reliability of international commitments." *American Journal of Political Science* 53(2):461–476.
- Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.
- Mattes, Michaela and Mariana Rodriguez. 2014. "Autocracies and international cooperation." *International Studies Quarterly* 58(3):527–538.

- Milner, Helen V. and Keiko Kubota. 2005. "Why the move to free trade? Democracy and trade policy in the developing countries." *International Organization* 59(1):107–143.
- Moon, Chungshik. 2015. "Foreign direct investment, commitment institutions, and time horizon: How some autocrats do better than others." *International Studies Quarterly* 59(2):344–356.
- Peceny, Mark, Caroline C. Beer and Shannon Sanchez-Terry. 2002. "Dictatorial Peace?" *American Political Science Review* 96(1):15–26.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.
- Smith, Benjamin. 2005. "Life of the party: The origins of regime breakdown and persistence under single-party rule." *World politics* 57(3):421–451.
- Way, Christopher and Jessica LP Weeks. 2014. "Making it personal: regime type and nuclear proliferation." *American Journal of Political Science* 58(3):705–719.
- Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106(2):326–347.
- Weeks, Jessica LP. 2014. *Dictators at war and peace*. Cornell University Press.
- Wilson, Matthew C. and James A. Piazza. 2013. "Autocracies and Terrorism: Conditioning Effects of Authoritarian Regime Type on Terrorist Attacks." *American Journal of Political Science* 57(4):941–955.