"Measuring What You Can't See"

The Latent Characteristics That Structure Autocratic Rule*

Barbara Geddes,† James Honaker,‡ and Joseph Wright§

August 20, 2014

Research in progress. Please do not distribute without permission.

Abstract

Research on autocratic regimes in comparative politics and international relations has burgeoned in the past two decades, and often uses categorical typologies of autocratic regimes to distinguish among different types of dictatorships. Building on advances in methods for estimating the latent dimensions of democracy, this paper uses historical data on 30 features of autocracies to estimate the latent dimensions of autocratic rule. We identify three time-varying dimensions of autocracy that correspond to the ideal types proposed in the original Geddes' typology: party dominance, military rule, and personalism. We show that the dimensions of autocratic rule are orthogonal to commonly-used measures of democracy-autocracy; compare these dimensions to existing typologies of autocratic regimes; and propose how this information can be structured as clusters. We show that while party dominance and military rule can be measured using existing data sets, the new time-varying measure of personalism is unique.

*Author order is alphabetical. This research is funded by the Minerva Research Initiative (ONR #N000141211004) and the National Science Foundation (BCS-0904478). We thank José Cheibub, Rob Franzese, Fabrice Lehoucq, Daniel Stegmueller, Milan Svolik, and participants at the University of Illinois Comparative Politics Workshop (Feburary 2014) for helpful feedback.

†UCLA. geddes@ucla.edu

‡Harvard University. jhonaker@iq.harvard.edu

§Pennsylvania State University. josephgwright@gmail.com (corresponding author).

> We learn from the Soviet experience that we have to strengthen the grip of the party on the military. The Soviet army de-politicised itself, dissociated itself from the party, and changed its nature (from a party army) to a national army. This means stripping the party of the ability to defend itself. – Chinese President Xi Jinping[1]

As many scholars know well, dictatorship in the Soviet Union under the Communist party differed markedly from the rule of the Chinese Communist Party (CCP). While both regimes ruled with a single, dominant political party, successfully institutionalized leadership succession, and lasted for over six decades, these dictatorships differed on other important dimensions, including the relationship between the party and the military. As the above quote illustrates, not only does the current leader of the CCP recognize these differences between the two Communist regimes but he suggests that these differences may change over time, with implications for how long autocratic regimes remain in power.

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 military conflict (Peceny, Beer and Sanchez-Terry 2002; Weeks 2012), civil war (Fjelde 2010; Gurses and Mason 2010), trade (Milner and Kubota 2005), terrorism (Aksoy, Carter and Wright 2012; Wilson and Piazza 2014; Conrad, Conrad and Young 2014), democratization (Geddes 1999; Brownlee 2009), economic growth (Gandhi 2008; Wright 2008), and domestic investment (Wright 2008; Gehlbach and Keefer 2012) demonstrate that variation within the group of countries categorized as non-democracies can provide substantial leverage on explaining many important outcomes. As this research grows, measures of many features of autocratic regimes – including typologies – have proliferated. This study uses recently coded data in a first attempt to identify the latent dimensions of autocratic rule.[2]

This paper makes two contributions to the literature on comparative authoritarianism. First, we establish that detailed historical data on the relationships between the leader, the party, and the

[1]Quoted in Ching Cheong, "Dreams of reform remain just that in China." *The Straits Times* 16 February 2013. Available at: `http://tinyurl.com/mu2xjuy`. [accessed 4 June 2013]

[2]We stress that this coding project is ongoing. The results reported in this draft are therefore preliminary.

military in autocratic contexts provides information that is largely orthogonal to most commonly used measures of the level of democraticness. This suggests that dictatorships differ from each other on dimensions that cannot be measured using democracy variables such as the Polity score. While many scholars implicitly make the same argument when they use information on autocratic regime types, this paper provides the first (to our knowledge) evidence to confirm this supposition.

Second, this paper uses a new approach to studying autocracies that makes use of recently coded data to construct the latent dimensions of autocratic rule. This approach differs from the existing practice of creating new categorical typologies of autocratic regimes. Rather, we aggregate existing information to understand the extent to which different variables – which may measure distinct concepts – capture variation in autocracies along distinct dimensions.

The Geddes (1999) typology has existed since the late 1990s. In the past 15 years scholars have introduced new typologies to measure additional concepts (Bueno de Mesquita et al. 2003; Hadenius and Teorell 2007; Gandhi 2008; Weeks 2012) and to circumscribe the universe of cases for particular research questions (Howard and Roessler 2006; Levitsky and Way 2010). Many studies also measure important features of autocratic rule that do not necessarily entail typologies but rather measure ordinal concepts. For example, Gandhi (2008) uses information on political parties and legislatures to construct an ordinal measure of institutionalization in dictatorships, which she then shows is associated with many important outcomes.

The present paper uses recently coded, time-varying data on 30 variables to construct measures of three latent dimensions of autocratic rule theorized in earlier work. We use an exploratory principal components approach that can incorporate information from other data sets – which potentially measure additional concepts along differentm, untheorized dimensions – to examine the extent to which these other measures overlap with the latent dimensions we construct or whether they capture new dimensions. With the number of existing measures of autocratic rule proliferating, we believe a next step is to examine how the existing data can be most appropriately structured for use in applied research.

The first section reviews different ways researchers measure important concepts in autocratic contexts, including regime typologies, formal political institutions, and combinations of these. The

second section discusses the growing literature on structuring data on democracy, including attempts to construct measures of the latent dimensions of democracy. The third section builds on these studies by using information on the relationships between autocratic leaders, their support parties, and their militaries to construct three variables that capture latent dimensions of autocratic rule that previous research using typologies has suggested influence the policies and international behavior of dictatorships. The fourth section demonstrates that these dimensions are not correlated with most extant measures of democracy-autocracy, while the fifth section compares them to existing classifications of autocratic rule introduced by Weeks (2012) and Geddes, Wright and Frantz (2014). We then show the latent measures we construct perform as expected in an empirical model of autocratic regime breakdown. The final section concludes with suggestions for future research.

Prior Research

Varieties of autocratic rule

Scholars of dictatorships have offered several categorizations of autocratic regimes (Huntington 1968; Wintrobe 1990; Huntington 1991; Bratton and van de Walle 1997; Chehabi and Linz 1998; Geddes 1999; Hadenius and Teorell 2007; Cheibub, Gandhi and Vreeland 2010). During the height of the third wave of democratization in the 1990s, comparative research turned to studying these transitions, which in turn spurred even more regime categorizations focusing on hybrid or transitional autocracies – countries that had the trappings of democracy, but where incumbent leaders and parties did not leave office in fair and free elections or where political participation was still severely restricted (Karl 1995; Diamond 2002; Schedler 2009; Levitsky and Way 2010). POur study builds on a categorization of dictatorships generated from the question: 'who rules constrained by whom?' (Brooker 2000). This approach draws basic insights from the sociological literature on military institutions and party organizations. It treats the interests and preferences of individuals and organized groups, such as the military or ruling party, that are able to influence autocratic decision-making as central to the task of distinguishing *types* (Janowitz 1960; Weber 1964; Huntington 1968; Nordlinger 1977; Linz 2000).

Other autocratic typologies, such as Hadenius and Teorell (2007), build on this sociological

approach by adding information on the formal institutional landscape of the regime. Their typology, for example, includes information on military rule as well as whether opposition parties are present in the legislature, thus allowing for institutionally distinct forms of military rule. Other typologies, such as Cheibub, Gandhi and Vreeland (2010), rely on information about the identity of the nominal regime leader – whether he is a monarch, soldier, or civilian – to classify autocracies.

In contrast to typologies based on a set of sociological characteristics or leader identity, still other studies focus on a specific attribute of autocratic rule to distinguish one group of autocracies from another or to assess the influence of a concept measured with ordinal values. For example, Gandhi (2008) theorizes the consequences of the level of institutionalization and uses leadership by military officers or monarchies as indicators of the organizational structures upon which a dictator can rely for support. Bueno de Mesquita et al. (2003) combine information on whether the nominal leader is a soldier with information from components of the Polity index to measure the size of the regime's support coalition and the size of the group that is eligible to select the leader. The degree of institutionalization and coalition size are ordinal meaures. Other studies pick a particular characteristic of autocratic rule, such as semi-competitive elections, to define the group of countries under study (Howard and Roessler 2006; Levitsky and Way 2010; Bunce and Wolchik 2011).

Our starting point for identifying the structure of autocratic rule is the Geddes' typology.[3] The classification of autocracies identifies individual *autocratic regimes*. A *regime* is defined as a set of formal and informal rules for choosing leaders and policies. The rule central to distinguish one autocratic regime from another in the same country is the rule that identifies the group from which leaders can be chosen and determines who influences leadership choice and policy. To remain in power, regime leaders must retain the support of members of this group, but leaders also have substantial ability to influence the membership of the group, especially after initial leadership selection. Autocratic *regimes* differ from autocratic *spells* (periods of uninterrupted non-democratic rule) and the tenure of individual autocratic *leaders*. For example, the post-1979 clerical regime in Iran is distinct from the pre-revolutionary regime under the Shah. These two regime together

[3]While we start with this typology, we emphasize that information from measures from other data sets can easily be incorporated into the analysis to assess whether these variables add information on other dimensions not captured in our initial dimensions.

constitute one autocratic spell because there was no democratic interlude between them. The clerical regime has had multiple leaders, Ruhollah Khomeini and his successor Ali Khamenei.

The original classification used a series of questions deemed theoretically relevant to understanding politics in autocracies, largely following historical studies of partiucular cases. The questions, which appear in the Appendix to Geddes (2003), are grouped into three categories to reflect three ideal types of autocracy: military, personalist, and party-based.[4] The original coding placed regimes, that is one continuous span of country-years coded as guided by the same basic rules, in one category or another based on whether there were a relatively high number of affirmative answers to questions within a particular category (Geddes 2003, 225). For example, if a regime received a high number of 'Yes' answers to questions pertaining to personalist dictatorships but a low number on questions addressing military and party rule, then the regime was classified as a personalist dictatorship. Hybrid regimes were those that scored relatively highly in more than one category of questions.

This approach to classifying regimes, while useful for many purposes, faces three issues which stem from the fact that the classification of regimes into exclusive categories reduces potentially relevant information.[5] First, using a relative cut-point on an index to delineate whether an individual regime falls into a particular category means that some concepts, which may vary in degree across all dictatorships, are reduced to a binary categorization. The original classification was aware of this issue and addressed it by allowing for hybrid categories. However this approach still obscures potentially relevant information.

Second, this method of aggregating information does not allow researchers to pinpoint the particular questions (and thus the concepts) that are most important for classifying regimes in one category and not another. For example, if a dictatorship is coded as party-based we do not know whether that is due to factors related to the organizational structure of the support party, to rules regulating leader succession, or to both equally. The inability to distinguish different concepts that are used in the original classification of exclusive categories may lead to *concept stretching*.

[4]Wright (2008) and the most recently updated version in Geddes, Wright and Frantz (2014) add monarchies and oligarchies.

[5]Hadenius and Teorell (2007), Svolik (2012), and Wilson (2014) discuss some of these issues.

For example, researchers have used the typology (or variations on it) to measure constraints on the leader (Weeks 2008), the breadth and depth of the support coalition (Wright 2009), and the range of available coercive and co-opting strategies (Wilson and Piazza 2014). These studies argue that the relevant concept is captured in some of the original questions, but the exclusive categories may be measuring other important concepts as well. Using the original classification of exclusive categories, we do not know the relative weight of particular concepts used to place a regime in one category or another.

Finally, the original typology placed a specific regime in the same category throughout its entire duration. Some of the information used to classify regimes, however, varies over time within the same regime. For example, the answers to questions that assess the extent to which the cabinet is comprised of civilians and whether the leader's relatives occupy senior military or party positions, can vary over time within the same regime.

The recently coded data, combined with our approach, addresses each of these issues by using the raw variables – some of which now vary over time within regimes – to structure the information into its principal components. This approach does not use abritrary cut-points in an index (except to assess the number of principal components); it allows us to see which raw variables contribute the most information to each dimension; and it allows for information on multiple dimensions to vary over time within a particular regime.

Latent dimensions of democracy

Characterizing and scaling the level of democracy has been a longstanding enterprise in political science. The depth of this literature and its importance to many questions might suggest that the measurement of regime type is a closed problem. Recent advances in this literature are largely concerned with aggregating the large array of measurements already in existence, rather than providing new theoretical constructs to measure.[6] These approaches rely on the multiplicity of past measurement approaches. We briefly summarize the principles behind the techniques used to aggregate existing measures of democracy, and then contrast them with the problems faced in

[6]However, see Coppedge et al. (2011) and Linzer and Staton (2011) for recent theoretical advances.

understanding the structure of autocratic rule.

The work of Pemstein and Melton (2010) assumes that alternate democraticness indices attempt to measure the same underlying concept, but arrive at different answers for some countries because of measurement error. This perspective assumes each variable is a proxy, generated by combining the true latent value of democracy with some coder-specific noise. To the extent alternate measures of democracy covary, the model assumes they are picking up the true latent value, and, to the extent they differ, the authors assume they are revealing the amount of measurement error in each index. The estimation goal is to reveal the underlying latent variable from the distribution of proxies. This approach relies on the existence of a number of high-level, high-structure variables coding the same quantity of interest, the level of democraticness.

In contrast, the work in the tradition of Item Response Theory (IRT), such as Treier and Jackman (2008), assumes that the constituent variables in any democracy index are low-level attributes of democratic states, but not individually sufficient measures of democracy. The IRT approach requires a large number of measures of attributes of democracy, which are conceptualized as individual tests that strong democracies would pass. There is a *correct* answer that easily aligns with increased levels of democracy. The IRT model measures the degree of latent democratic *ability* by estimating both how difficult it is to get the correct democratic response on each attribute, and thus how well that measure discriminates the level of democracy, and how many of the these test measures are correctly answered or passed. This approach, however, assumes a large number of low-level, low-structure features of democratic regimes, each of which we understand as contributing to the same latent dimension in one direction. If dictatorships are multidimensional, and if we do not know *a priori* which questions scale on which dimension, or even in which direction the "correct" answer points, we can not implement this approach or monitor for overfitting.

The Structure of Autocratic Rule

Little of what we want to understand about autocratic behavior can be explained by measures of "lack of democracy". Empirical research on autocracies shows that other basic characteristics of autocratic structure have real-world consequences. The features differentiating autocratic regimes

from each other may in fact be orthogonal to the characteristics measured by democraticness scales. The latent variable understanding, and techniques, however, can also be used to measure empirically salient autocratic characteristics. We continue in the latent variable tradition: for each regime-year there is a true underlying, but unobserved, location in the space of all possible regime dimensions. However, in the absence of either a large set of uni-directional sub-measures along a single dimension, or a literature of previous dictatorship scalings, we depart from the accepted IRT and measurement error models.

While there are an increasing number of categorical and ordered measures of autocracy, as noted above, they frequently mix distinct underlying concepts, such as the breath of the support coalition (Bueno de Mesquita et al. 2003), the extent of multiparty competition (Hadenius and Teorell 2007) and an individual trait of the nominal leader (Cheibub, Gandhi and Vreeland 2010). Thus we do not have multiple measures of the same concept in the extant literature that can be combined in a measurement error approach.[7] Consequently, we start with more flexible, exploratory methods to measure autocratic structure.

Figure 1 depicts our conceptualization of the measure space of state structure: there is a dimension, d that measures democraticness. As countries become more democratic (moving to the right), they become increasingly alike. As states become less democratic, they begin to have other dimensions that define their autocratic "type" (here, these dimensions are labeled a_1 and a_2, although the dimensionality could be greater). These dimensions are orthogonal to the measure of democraticness. That is, knowing the locations in a_1 or a_2 would not contribute to knowing the location in d. Our estimation goal is thus to reveal both the underlying dimensions of authoritarian structure and the location of the latent position of all regimes in these dimensions.

[7]For some concepts, such as military rule or party strength, there may be enough distinct data sets to employ a measurement error approach. However, some of the autocratic data sets, such as Hadenius and Teorell (2007), already combine information from other existing data sets, such as the Freedom House index, so it may not be appropriate to treat each individual data set as an independent measure.

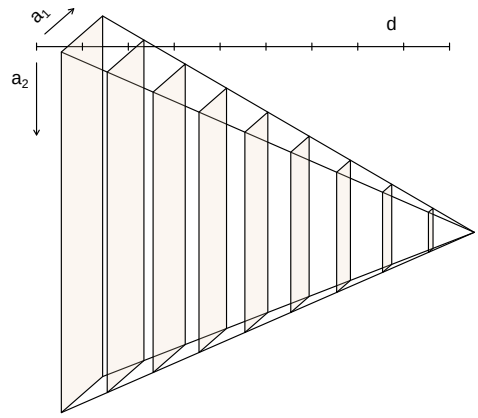


Figure 1: A model of our conceptualization of the space of regimes. A unidimensional level of democraticness, d, makes states more alike as they increase in democraticness (moving to the right in the space above), while increasingly undemocratic states differ along other dimensions of authoritarianism, here a_1, a_2. The dimensions of a are orthogonal to democraticness, and do not influence the measures of democracy in the literature, but do determine the ways in which authoritarian regimes maintain power and fail.

Latent dimensions of autocratic rule

In this section we describe how we aggregate information on the structure of autocratic rule. We use information from the time-varying coding of some of the questions used to create the orginal Geddes' typology. The raw data contain information from 30 variables constructed from questions about the relationships between the dictator, the political party that supports the regime (if there is one), and the military.[8] The variable names, their value definitions, and summary statistics are in the Appendix.

We stress that the coding of these variables is still in progress so the results we report should be interpreted with appropriate caution. Because the coding is still in progress, missing values are present. The number of missing values, and their patterns across variables, are described in the appendix. A limitation of our principal component analysis – common to most statistical models although not always present in data reduction methods – is that it requires a *rectangular* dataset, that is, all observations for all variables that are included. To run principal components then, we

[8]Coders are asked to record information for January 1 for each calendar year.

multiply imputed all the missing values (Honaker, King and Blackwell 2011).

Multiple imputation is a well accepted and increasingly common approach to missing data problems in statistical analysis (Rubin 1976; Schafer 1997; King et al. 2001). The common focus of explanations of multiple imputation is often about recovering missing information, because of the ability of imputation models to use auxiliary forecasting variables in the imputation model that are not present in the analysis model. However, the key objective of multiple imputation, especially in our setting where no extra variables are present, is to create a complete dataset where all the sufficient statistics, and relationships between variables are exactly the same as in the original observed, but partially completed, dataset. Thus our goal is to fill in the missing data while not changing any relationship or information in the dataset, so that we are able to use principal components to summarize these relationships. Many of these variables are static or move slowly, within regimes, while likely abruptly changing across regimes. We used a model to capture the patterns across time, by allowing the imputation model to have a pattern across time, in any variable, within any regime (Honaker and King 2010), although we found this did not give us appreciably different results than simpler imputation models.

The plots in Figure 2 show factors in a three-dimensional space derived from constructing the first three dimnesions.[9] Each panel shows two dimensions plotted against each other. Each circle in the plots is a separate country-year observation, while the arrows represent each variable employed in the factor analysis. The direction of each arrow shows how it contributes to the two dimensions represented along each axis and the length of each arrow describes how much weight each variable contributes to the factor loading. Arrows 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 initial 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 diffilcult to distinguish visually, the cluster of arrows pointing East contain information on this concept: e.g., whether a support party exists (`supportparty`); whether the

[9]Figure A-2 shows the eigen values for all 30 dimensions. For this analysis, we only examine the first three, which capture 80 percent of the variation.

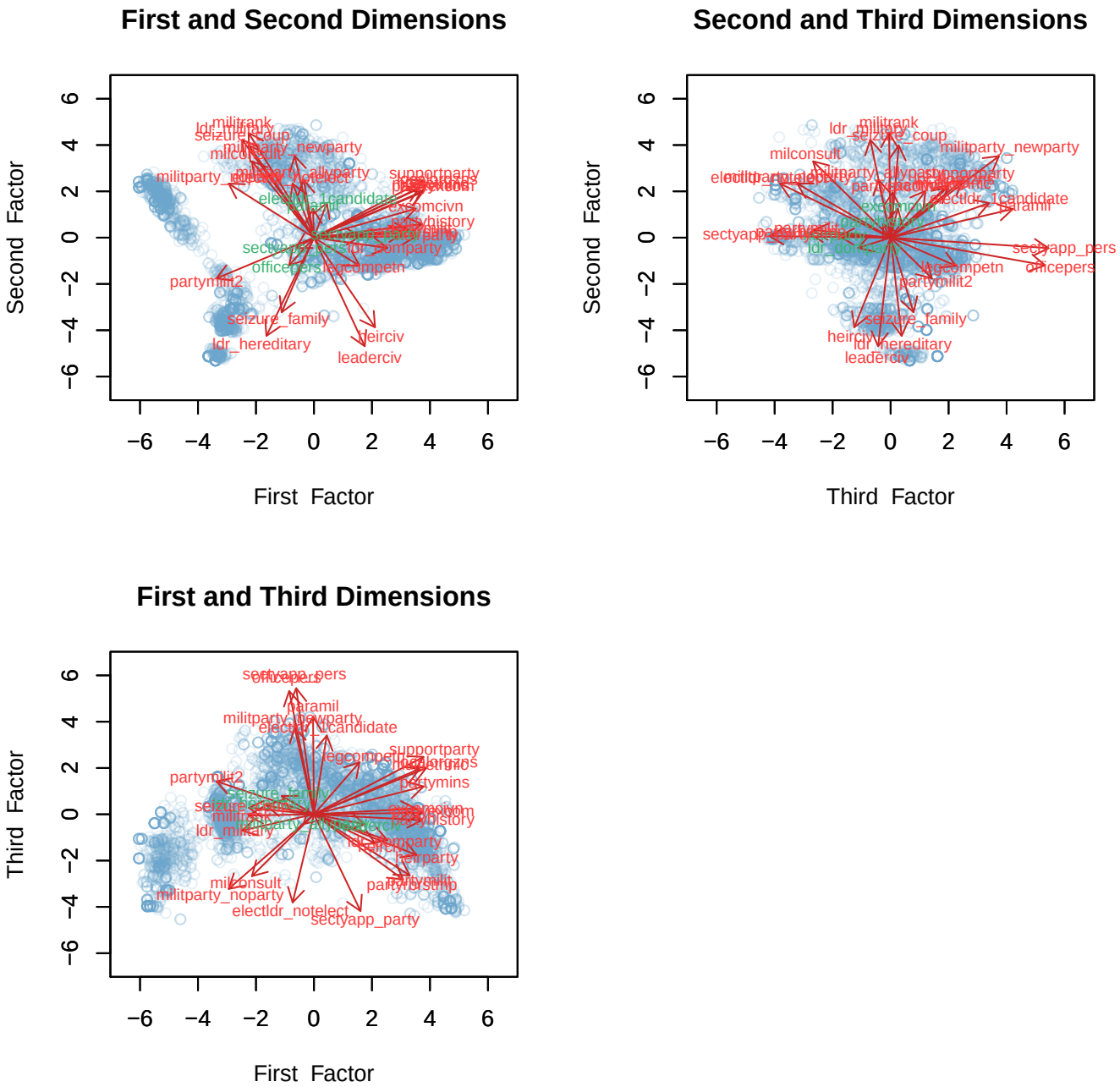


Figure 2: Three components of autocratic rule.

support party has local-level branch organizations that link party militants to citizens (`localorgs`); and whether the regime leader was selected from the dominant party (`ldr_domparty`). The second dimension, displayed on the vertical axis in the two upper plots, 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 (`ldr_military`); whether the regime leader was a high ranking military officer prior to assuming power (`militrank`); and whether there is a routine mechanism for the leader to consult the military in policy decisions (`milconsult`). The third dimension, shown on the horizontal 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` aand `sectyapp_party`) and whether the dictator personally controls appointment to high office (`officepers`).

Table 1 shows the individual variables from the raw data that contribute the most information to each of the first three dimensions.[10] The small colored diagrams next to each variable name describe the information each variable contributes to all three dimensions. For example, `support_party` and `localorgs` are the variables that contribute the most information to the first dimension, as represented by the red line in the diagram next to the name for each variable. These two variables also contribute some information to the second and third dimensions, as represented by the shorter blue (second dimension) and green (third dimension) lines in each diagram. This pattern reflects the real world in which many personalist and military regimes have support parties. The variable that distinguishes most among the dimensions is `partyhistory`. Strong, party-based regimes are most often supported by parties that were originally organized to lead revolutions or run in competitive elections, whiel the parties that support personalisst regimes were often created after the seizure of power. Some military dictatorships also create parties after seizure, but others ally with pre-existing parties. Inspection of variables that contribute the most information to the first dimension again suggest this dimension captures the strength of the party.

The variables listed in the second column are those that contribute the most information to

[10]Table A-2 lists the variables in the factor analysis as well as the factor loading for the first three dimensions.

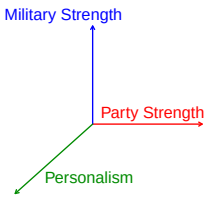
	First Dimension Party Strength	Second Dimension Military Strength	Third Dimension Personalism
1	localorgzns	leaderciv	sectyapp_pers
2	supportparty	militrank	officepers
3	partyhistory	ldr_military	paramil
4	partymins	ldr_hereditary	sectyapp_party
5	partyexcom	seizure coup	militparty_newparty
6	multiethnic	heirciv	electldr_notelect
7	excomcivn	milconsult	electldr_1candidate
8	heirparty	seizure_family	militparty_noparty
9	partymilit	militparty_allyparty	legcompetn
10	partymilit2		
11	partyrbrstmp		
12	ldr_domparty		

Table 1: Variables arranged by loading on principal components. Variables are sorted between columns based on which dimension of the principal components they contribute the most to. Within columns, variables are sorted by magnitude, with the variables with the largest factors in that dimension at the top. Variable definitions are in the Appendix.

the second dimension. For example, the two variables that contribute the most information are `leaderciv` and `militrank`. The first measures whether the regime leader was a member of the military prior to taking office and the second his military rank (e.g. general or colonel) if he was a member of the military prior to assuming power. Regimes based on the military as an institution are almost always led by highly ranked officers, not sergeants or majors. In constrast to the first column, the variables in the second column suggest this dimension measures characteristics related to the military, including adherence to military norms about hierarchy.

Finally, variables in the third column, such as `sectapp_pers` – which measures whether the regime leader personally controls key organizations in the security apparatus – and `officepers` – which measures whether the leader has descretion over appointments to high office – tap into concepts related to the personal power of the regime leader. The regime leader, in the Geddes' conceptualization of regimes, has more power when he has control over appointments to high office and controls a separate security or paramilitary organization to counter the power of the pre-existing military institution, especially the latter's capacity to credibly threaten a successful coup.

Measuring autocratic rule in China

As Xi Jinping noted in his 2012 speech, autocratic rule in China has changed since the end of the civil war in 1949. However, the CCP is coded as a dominant or one-party civilian regime during its entire time in power by standard autocratic typologies such as Geddes (1999), Hadenius and Teorell (2007), and Cheibub, Gandhi and Vreeland (2010). Figure 3 shows the changes in party and personalist dimensions in China over six decades. In particular, the figure shows the rise and fall of Mao's personal power, which is partly reflected in a decrease in party power in the later 1960s and subsequent rise in the 1970s.

After the civil war, the Communist Party (CCP) elite ruled China; and although Mao's power was rising relative to others in the party's inner circle, "decisions were usually made by a small group whose composition for over 30 years (until mid-1966) was amazingly stable" (Robinson 1972, 157). Between 1959 and 1961, moderates led by Liu Shaoqi and Deng Xiaoping, whose support came especially from the party machine, gained sway while Mao lost influence (Shinn and

L.Worden 1987). In mid-1966, however, Mao launched the Cultural Revolution to attack the party bureaucracy and undermine moderate members of the top leadership like Liu and Deng (Israel 1974, 405). Mao used mass rallies at which crowds responded "ecstatically" to mobilize popular support for the campaign and to "intimidate opponents" within party (Israel 1974, 412, 423). The military, commanded by a Mao protégé, provided transport to rallies and support for Red Guard volunteers as they fanned out around the country (Israel 1974, 414). The leaders who had been most involved in building the party were attacked by Red Guards and purged during the Cultural Revolution (Klein and Hager 1974, 225), as were state and party bureaucracies, which weakened the party's hold on the state apparatus and the country. The Central Committee Secretariat, the party's administrative apparatus, was abolished (Brooker 1995, 93-94). At the 1969 CCP Congress, the purged party stalwarts were replaced on the Central Committee and Politburo by military officers and other Mao loyalists, including his wife – the peak of Mao's personal dominance over the party (Klein and Hager 1974, 224-26, 239; Scalapino 1972, 96-98).

By the early 1970s, the party had begun to reassert itself. After Mao died in 1976, members of the Politburo quickly arrested Mao's wife and the rest of the "Gang of Four" leaders of the anti-party faction to prevent a power struggle. In 1977, the Central Committee exonerated Deng Xiaoping, and he resumed all posts. Factional struggles continued during the late 1970s as Deng gradually reasserted his and party pragmatists' dominance (Shinn and L.Worden 1987). In 1980 the Party Secretariat was reconstituted (Ristaino 1987). The post of party chair was abolished in 1981 in favor of a more collegial senior leadership (Brooker 1997, 25). Since Deng's death, the party has enforced term limits for top leaders and a somewhat collegial leadership.

This example illustrates that, with the appropriate data, it is possible not only to measure the level of personalism as it varies over time within a particular regime – for example the six plus decades of Communist rule in China – but that we can measure features of authoritarianism as they vary over time within the rule of an individual leader, such as the rise and fall of Mao's power. This feature of the data improves on extant typologies, such as Cheibub, Gandhi and Vreeland (2010) and Weeks (2012), that capture some of the variation between leaders within a particular regime, but not changes over time within the tenure of leaders. Indeed, central to understanding

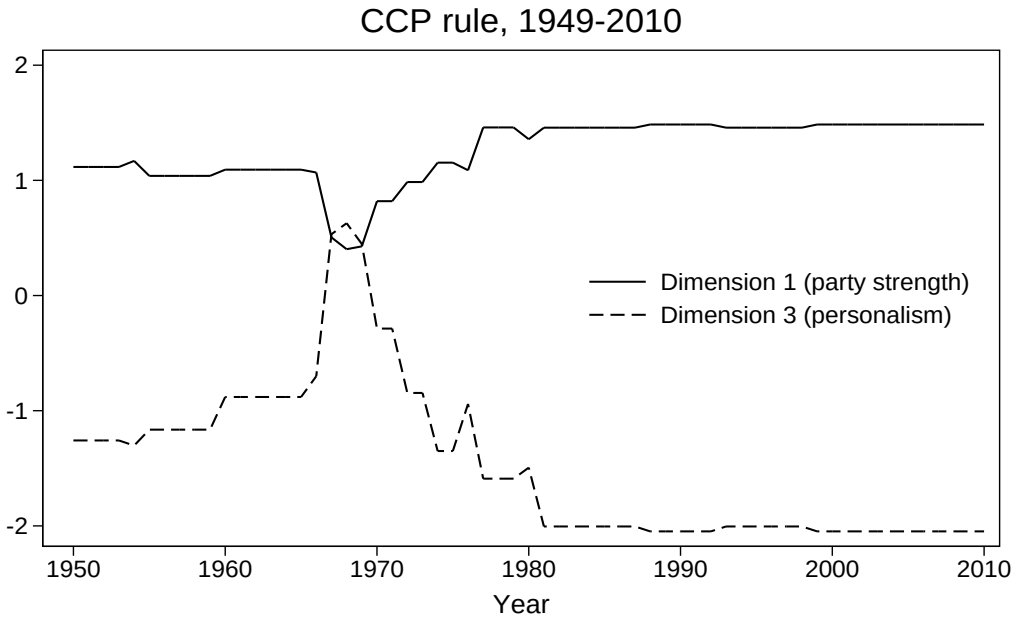


Figure 3: Two dimensions of autocratic rule in China.

personalist rule is explaining *when* an individual leader consolidates personalist power.

Latent dimensions of autocratic rule and measures of democracy

In this section we show that the dimensions of autocratic rule are *not* correlated with most commonly-used continuous measures of democracy. Figure 4 shows a correlation matrix between three dimensions of autocratic rule and extant democracy variables. Each cell in this matrix is color coded to more easily discern patterns in the correlations, with cells that are more red denoting greater correlations. We can see several things. First, the democracy scores all highly correlate with each other, forming a distinct red block in the matrix. The only exception to this is Coppedge's inclusiveness measure. Inclusiveness does not correlate with any other democracy measure above 0.35, while no other pair of democracy measures correlate below 0.52, with a mean of 0.75.

Second, none of the three dimensions of autocratic rule correlate with common cross-national measures of democracy, including components of the Polity score, the Freedom House measure, and the Unified Democracy Score. The only exception is the first dimension, which we believe measures the strength of the support party. It is correlated with *inclusiveness* at 0.52. The

concepts captured in the latter include an index of participation (Vanhanen) as well as measures of adult suffrage and women's political rights (Coppedge, Alvarez and Maldonado 2008). The correlation reflects the universal suffrage and high participation rates characteristic of dominant-party dictatorships. Overall, however, the bivariate correlations suggest that the dimensions of autocratic rule – especially the strength of the military institution and personalism – cannot be easily measured with existing data.

	latentdim1	latentdim2	latentdim3	polity2	exrec	exconst	udsscore	inclusiveness	contestation	freedomhouse
latentdim1	1	0	0	0.14	0.16	0.3	0.11	0.57	0.12	0
latentdim2	0	1	0	0.02	0.26	−0.11	0.06	0.03	−0.03	−0.12
latentdim3	0	0	1	−0.11	−0.1	−0.17	−0.02	0.15	−0.07	−0.02
polity2	0.14	0.02	−0.11	1	0.82	0.82	0.82	0.18	0.79	0.67
exrec	0.16	0.26	−0.1	0.82	1	0.55	0.7	0.15	0.64	0.52
exconst	0.3	−0.11	−0.17	0.82	0.55	1	0.72	0.35	0.78	0.61
udsscore	0.11	0.06	−0.02	0.82	0.7	0.72	1	0.34	0.87	0.87
inclusiveness	0.57	0.03	0.15	0.18	0.15	0.35	0.34	1	0.35	0.23
contestation	0.12	−0.03	−0.07	0.79	0.64	0.78	0.87	0.35	1	0.87
freedomhouse	0	−0.12	−0.02	0.67	0.52	0.61	0.87	0.23	0.87	1

Figure 4: A correlation matrix of our three latent dimensions of autocratic structure, and several measures of democracy. Visually, cells that are more red, show higher correlation. Democracy scores largely group together as a high correlation block, while our latent dimensions are not in that block.

In the Appendix, we provide further evidence that the three dimensions do not correlate with most measures of democracy. Figure A-3 displays scatterplots of latent autocratic dimensions with various measures of democracy, as well as the bivariate regression lines and spline smoothers. It again demonstrates that none of the latent dimensions are strongly correlated with any of the democracy measures, with the exception of a positive relationship between Coppedge's *inclusiveness*

measure and the first dimension (party strength).

We also provide a detailed comparision between the latent dimensions and the most commonly used democracy measure from the Polity project. Figure A-4 breaks down our representation of the latent space, visualized previously in Figure 2, by each level of the Polity index. It shows that observations in the latent authoritarian space do not cluster together at different levels of Polity, or otherwise move systematically. Only at the extreme bound, with a Polity score of -10, is there a perceived pattern to the location of the authoritarian states in our dimensions at any level of Polity. A Polity score of -10 then, predicts the cluster of monarchies – because of specific Polity coding rules – but does not otherwise capture authoritarian structure. To demonstrate this point further, Figure A-4 presents the means and standard deviations of our scores across the range of the Polity index to show that in the range -9 to 0 there is no discernable relationship between our latent positions and the Polity score. However, as the Polity score increases, indicating greater democraticness, the variances are decreasing, suggesting that autocratic states at high levels of Polity are becoming more alike and more like democracies while the autocratic structure of politics is less distinguishable.

Latent dimensions of autocratic rule and regime typologies

In this section we compare the latent autocratic space to two closely related typologies of autocratic regimes, from Weeks (2012) and Geddes, Wright and Frantz (2014). The latter is an update of prior versions of the original Geddes' typology. The former builds on information from the original questions used to code the Geddes' typology.

To measure personal and military traits in a more nuanced way, Weeks (2012, 356) constructs indices from the raw data, using eight variables to measure personalism and five to measure military rule. From this she: (a) "create[s] indices representing the proportion of 'yes' answers" on the two dimensions"; and (b) "create[s] dummy variables for each of the four regime types, using a cutoff of 0.5 to classify countries as either personalist or nonpersonalist, or military or civilian." Combining information from these dummy variables, she creates four "types" of autocratic regimes: *strongman*, which is military and personalist; *junta* which is military but not personalist; *boss* which is not

military but personalist; and *machine* which is neither. She then adds two more types to complete the universe of autocracies: *monarchy* and *other*.

Variables	Dimension 1	Dimension 2	Dimension 3
Strongman (Weeks)	-0.221	0.466	0.184
Junta (Weeks)	-0.296	0.363	-0.134
Boss (Weeks)	0.154	-0.010	0.206
Machine (Weeks)	0.507	-0.083	-0.266
Personalist index (Weeks)	-0.174	0.057	0.455
Military index (Weeks)	-0.731	0.731	0.026
Military regime (GWF)	-0.472	0.412	-0.248
Monarchy (GWF)	-0.371	-0.678	0.065
Party regime (GWF)	0.727	0.079	-0.152
Personalist regime (GWF)	-0.200	0.155	0.331

Table 2: *Correlations of extant typologies with three dimensions of autocratic structure.*

Table 2 shows the bivariate correlations between the three dimensions and these extant types. The first panel shows the correlations with the main Weeks' categories. The first dimension is positively correlated with *Machine*, while the second is positively correlated with both *Strongman* and *Junta*. The third dimension is not strongly correlated with any of the Weeks categories. The next panel shows the correlation with the indices Weeks created. Dimension 1 is not correlated with either index, but the second dimension is strongly correlated with the military index while the third is correlated with the personalist index. Finally, the bottom panel shows that Dimension 1 is positively correlated with *Party* and negatively correlated with *Military* and *Monarchy*, while the second is positively correlated with *Military* and negatively with *Monarchy*; and the third dimension is positively correlated with *Personalist*.[11] These correlations confirm that the first dimension is picking up party strength, while second measures military strength, and the third personalism. Further, *Monarchy* is negatively correlated with the first two dimensions, suggesting that it might

[11]In this table, and the figures below, we collapse the hybrid categories into the pure types (e.g. military and military-personal are treated as military) to keep the number of categories manageable, though collapsing them in this was inevitably increases the error in classifications.

be described as a regime type with both weak parties and weak militaries.

Figure 5 plots two dimensions for each of the regime categories in Geddes, Wright and Frantz (2014). Each graph also displays a vertical line at -0.75 on the first dimension to mark a cut-point between the two modes of the distibution of this dimension. A horizontal line at zero shows a possible (arbitrary) cut-point along the military dimension. Nearly all the observations in the top left category, *Military*, fall above zero on the vertical axis, which we believe measures military strength. However, there are two clusters of observations that differ in the strength of the support party. Observations in the *Personalist* category fall in all quadrants of the two-dimensional space described by the first two dimensions, which suggests that these regimes vary in both military and party strength. The observations in the upper left panel, *Party* regimes, mostly lie to the right of -0.75, indicating stronger support parties. Finally, *Monarchies* cluster in the bottom, left corner of the space, suggesting both weak militaries and weak parties.

Figure 6 plots the data in the latent space described by the second and third dimensions, which we believe measure military strength and personalism. As expected, almost all the observations in the *Military* category lie to the right of zero on the horizontal axis measuring the second dimension. In the *Personalist* category, however, observations lie in all parts of the space. If these observations were correctly categorized as *Personalist*, we would expect them to lie above zero on the vertical dimension, measuring personalism. *Party* regimes lie in almost all parts of the space, while monarchies are located mostly in a cluster with low values in the second dimension (military strength) and near zero on the vertical axis (personalism). This latter plot suggests that while monarchies appear to have weak militaries, they vary in the extent of personalism.

These figures illustrates two points. First, even though *Military* and *Party* are observed where we would expect them, these figures show that there is substantial variation on other dimensions not used to categorized a particular regime. For example, though all *Military* regimes have high scores on dimension 2 (military strength), there is variation in both party strength and personalism among this group of observations. A similar point holds with respect to *Party regimes*. Thus, the dimensional analysis approach allows us to measure variation in the world of autocracies that cannot be captured in exclusive regime categories.

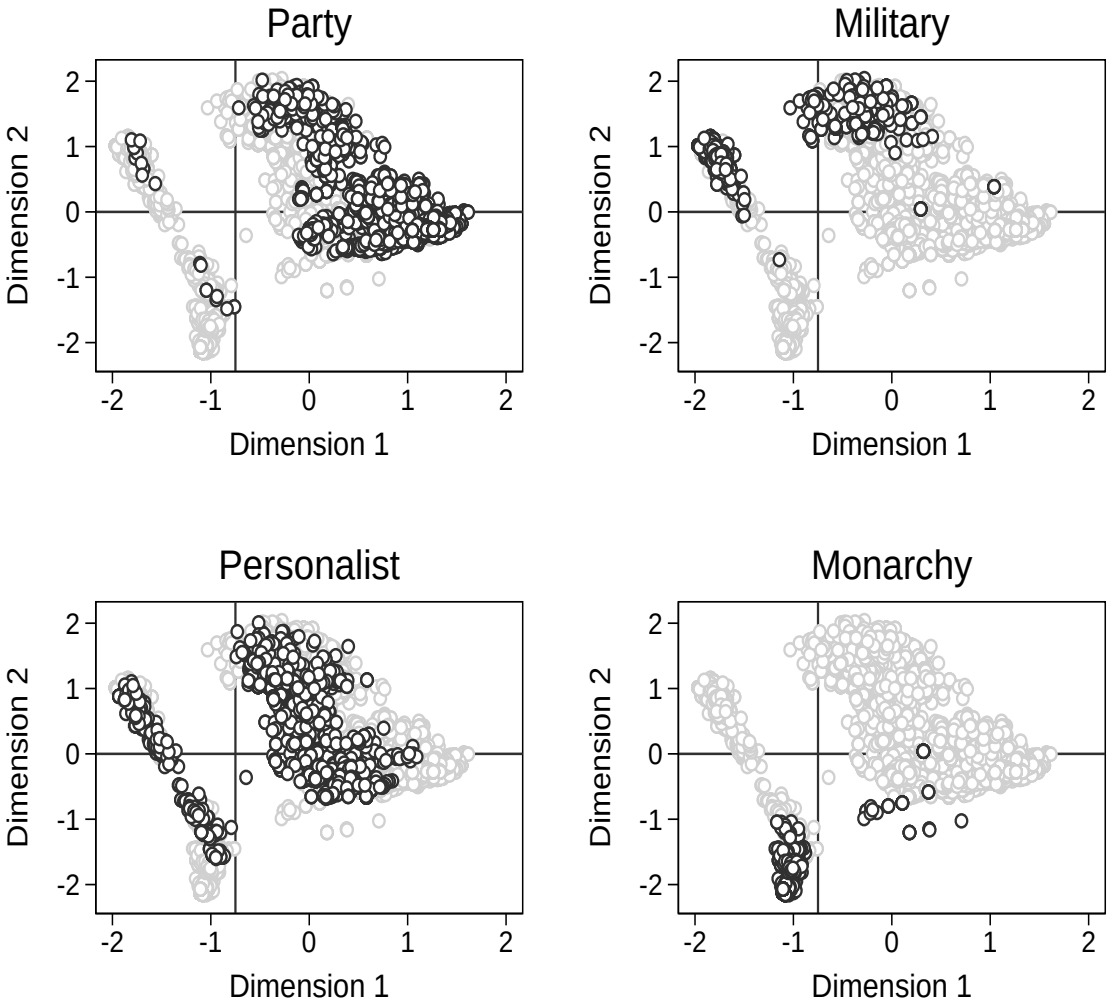


Figure 5: Party and military dimensions of autocratic rule, by Geddes, Wright, Frantz (2014) typology.

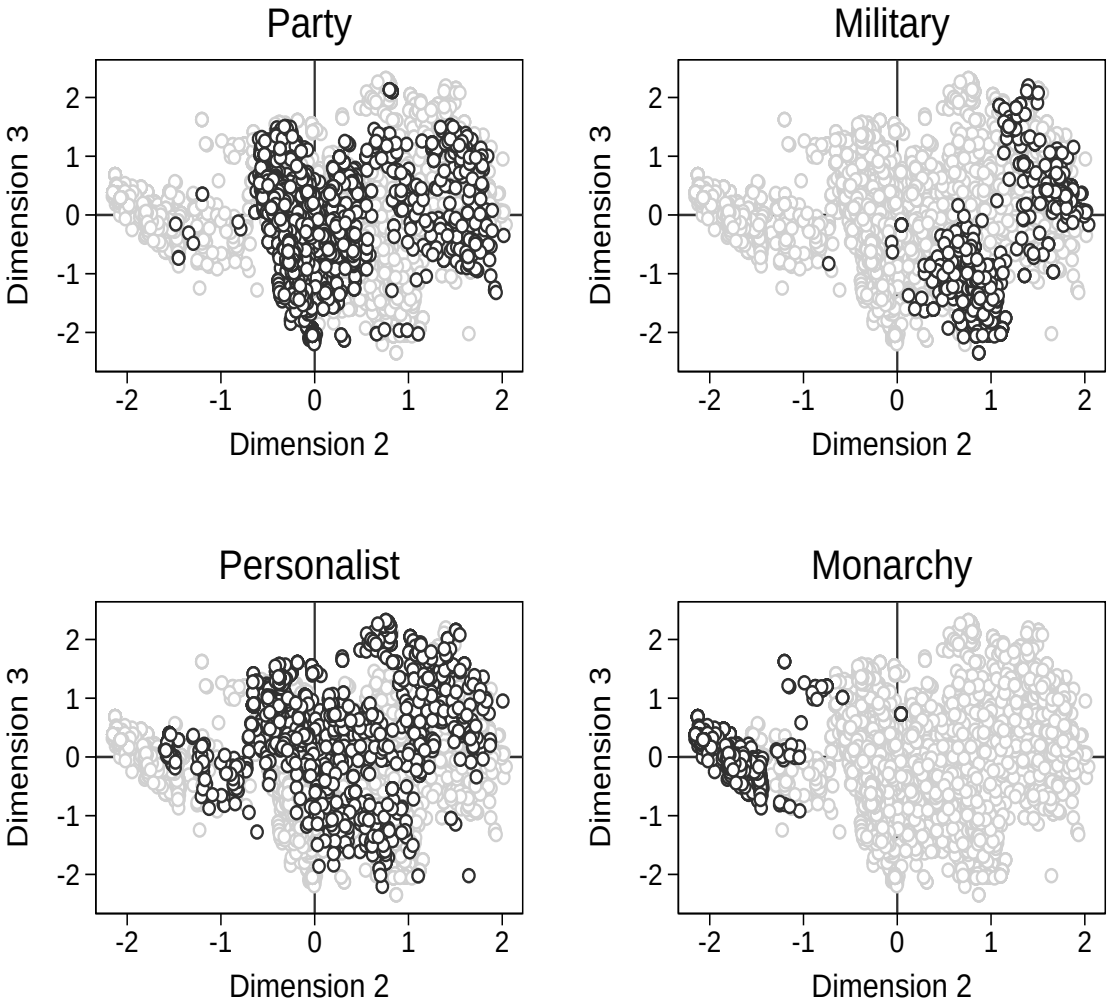


Figure 6: Military and personalist dimensions of autocratic rule, by Geddes, Wright, Frantz (2014) typology.

Second, the dimensions we identify from the raw data match relatively well with prior expectations about where observations categorized by Geddes, Wright and Frantz (2014) for *Military*, *Monarchy* and *Party* regimes should fall in the 3-dimensional space. However, the observations coded as *Personalist* regimes by Geddes, Wright and Frantz (2014) do not all have high scores on third dimension, which we believe measures personalism. Thus the party and military dimensions place individual country-year observations in the expected space for regimes placed in a particular category or type, except for those coded previously as personalist dictatorships – an anamoly we explain in the next section.

Figures 5 and 6 show that the latent dimensions fit the original Geddes' categories. In the Appendix, we show that the first two latent dimensions – which we believe capture party and military strength – are also correlated with other existing measures of these concepts. For example, the first dimension correlates at 0.50 or better with four independent measures of the exisitance of one or more parties from Beck et al. (2001), Hadenius and Teorell (2007), Cheibub, Gandhi and Vreeland (2010), and Svolik (2012); and the second dimension correlates at 0.60 or better with four measures of military regime from the same data sets. However, the third dimension – which we believe measures personalism – does not correlate at 0.25 or more with any institutional (legislature or party) or military variables from these data sets.

These comparisons with existing data sources suggest two points. First, information from existing data sets may improve measures of the first and second dimensions. This also means that these concepts can be measured using extant data. However, because the third dimension does not correlate with variables from existing data sets, we believe it captures a distinct concept that can be measured from information that is unique to this data.

Improving existing approaches

While the correlations presented in Table 2 show that the three dimensions in our latent space largely correspond to categories from closely related typologies, visual inspection of Figures 5 and 6 shows there may be some mis-classified observations. Further, both typologies resort to adhoc categories for difficult to code cases. The Weeks' typology includes a substantial number in the

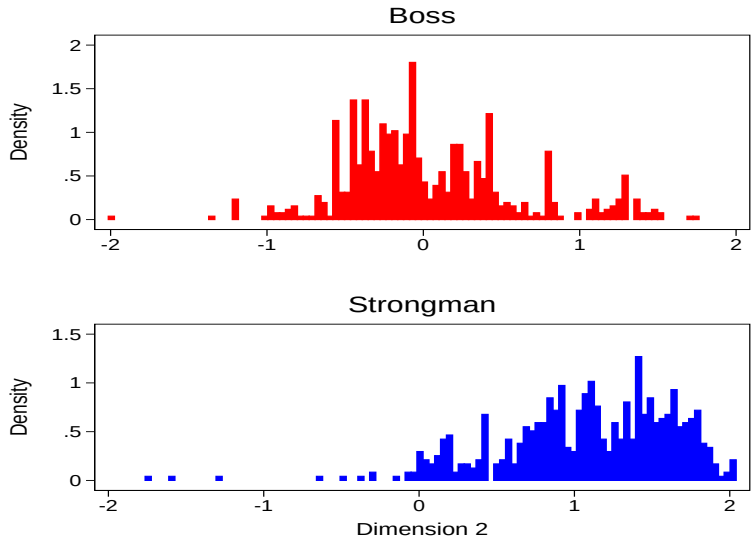


Figure 7: Distribution of second dimension in *Strongman* and *Boss* categories.

Other category and the Geddes' typology has hybrid regimes, with the latest update adding a new category (oligarchies) to classify some anomolous cases. The latent dimensions approach helps address each of these issues.

First, we flag some potentially mis-classified cases. Weeks classifies dictatorships using two concepts: personalism and militarism to construct four categories of regimes. *Boss* and *Strongman*, are two categories that should theoretically differ along one dimension (military) but not the other (personalism). Figure 7 shows that there are some observations in the *Boss* category that lie between 0 and 2 along the second dimension and thus share a similar level of military strength as the bulk of the observations in the *Strongman* category.

Second, the monarchy category in these typologies has some cases with relatively high values on first dimension, suggesting that these cases function differently than other monarchies. Examinging these cases suggests as much. For example, Burundi's first post-independence regime (until a 1966 coup) and the Shah's regime in Iran were monarchies, where leaders were chosen by hereditary succession rules. But each had a political party that supported the regime and a nontrivial share of the cabinet was from the party – features of party strength absent in most monarchies. The latent dimension approach identifies cases such as these where typologies may miss important aspects of

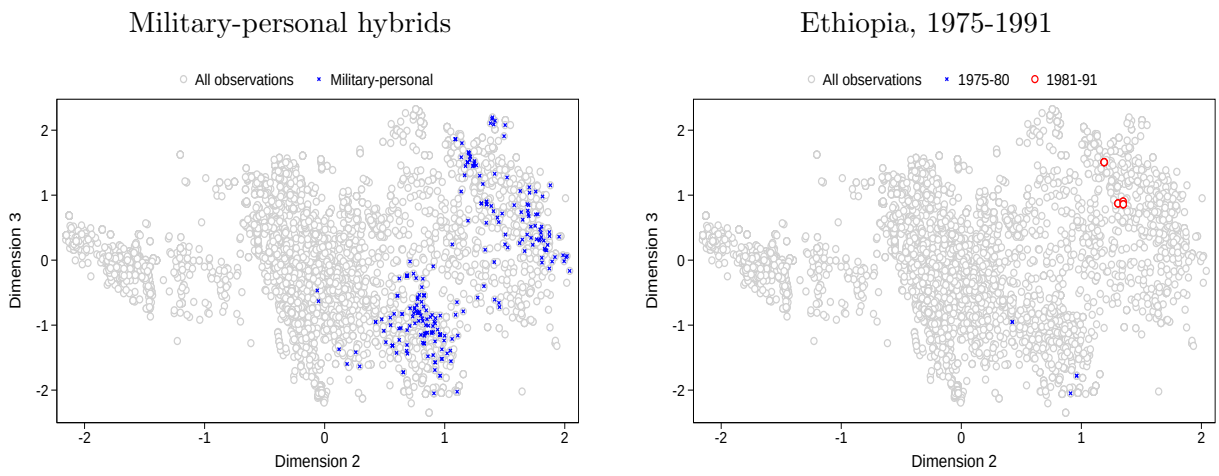


Figure 8: Military and personalist dimensions in Military-Personalist hybrid regimes.

party strength that are theoretically important in some applications.

Finally, the Geddes typology contains four hybrid categories, including one for regimes that have features of military, party, and personalist dictatorships. One approach to dealing with these cases is to absorb them into another category. For example, Wright (2009) codes all hybrid regimes with a party component as party regimes, but groups military-personalist hybrids with pure military regimes. The left plot in Figure 8 shows where military-personalist hybrid observations fall along two dimensions that we believe capture military strength (horizontal) and personalism (vertical). A prior expectation is that these observations, because they are military-personal hybrids, should cluster in the upper right corner. However, while they all fall above zero on the military dimension, they also fall in all areas along the personalist dimension.

One reason for this apparent anomaly may be that the level of personalism varies over time within particular regimes, as Geddes noted in the original explication of the coding scheme. In fact, she omitted the first three years of each regime from early regime classification data sets because so many military or party-led autocracies were becoming more personalistic during these first years in power. The variation in personalism over time is one of the central criticisms of the original Geddes' typology: personalism not only varies among regime categorized as non-personalist in the original but personalism may also vary within regimes as individual dictators consolidate greater personal power.

The right plot in Figure 8 shows where observations for one military-personalist hybrid regime, Mengistu's regime in Ethiopia from 1975-1991, lie in the space described by the second (military) and third (personalism) dimensions. The first six years of this regime are measured as relatively strong military features (i.e. above 0 on the horizontal axis) but relatvely weak on the personalist dimension (less than 0 on the vertical axis). The measure of personalism changes, however, by 1980, reflecting the centralization of power in Mengistu's hands. Initially, the dictatorship that ousted Emperor Haile Selassie in 1974 was governed by a committee made up of representatives of all military units (the Derg or PMAC). It was "controlled from below by the young officers of the army's various units" (Erlich 1983, 475). The absence of a strongman led to the "need for consensus on the major decisions" (Erlich 1983, 475). In 1977 Colonel Mengistu Haile Mariam defeated more moderate factions to become the third leader of the Derg and began consolidating personal control over appointments and decision making. By late 1979, he had "restructured the PMAC and filled all key positions both in the PMAC and in his government with loyalists" (Haile-Selassie 1997, 207). Then, to reduce the influence of all but a handful of officers, Mengistu initiated the Commission for Organizing the Workers' Party of Ethiopia (COPWE), which functioned as a highly disciplined party (and is coded as a regime support party beginning Jan 1, 1980), to counterbalance the military represented in PMAC (Haile-Selassie 1997, 232-33). Mengistu chaired the party, and party statutes gave him the right to appoint all members of its Central Committee, Executive Committee, and Secretariat as well issuing regulations for the admission of ordinary party members. Clapham (1988, 70) reports that Mengistu "spent many hours interviewing and selecting" the party's leadership cadre. With control over the composition of both party and PMAC leaderships, Mengistu had achieved substantial personal discretion over regime decision making by mid-1980. For the following decade, the regime is therefore measured as both highly personalist and with a high degree of military strength.

This example illustrates two improvements upon the exclusive categories. First, hybrid regimes can be characterized along multiple dimensions that may be theoretically important. Second, many hybrid regimes are coded as such because the strength of a particular dimension (or concept) varies over time within the regime's lifetime. For example, the level of personalism can increase when

an individual leaders successfully consolidates power. Again, the latent dimensions approach can account for these changes over time.

Constructing regime type categories with cluster analysis

While measuring the structure of autocratic rule as latent dimensions may be appropriate for many applications, some research designs might still benefit from employing exclusive categories of regime types. The raw data can also be structured as categories using cluster analysis. In this section, we provide one example from a k-median partition clustering algorithm. We chose a partitioning approach because we can specify a priori the number of clusters, 4 in our case. This number simply reflects an attempt to mirror the aggregated regime categories in Geddes, Wright and Frantz (2014): military, monarchy, party, and personalist.

Figure A-6 in the appendix shows how well clusters from various distance measures classify the regime type categories in Geddes, Wright and Frantz (2014).[12] For each set of clusters derived from a different distance measure, we conducted a ROC test using the clusters and the regime type categories (military, monarchy, party, personalist) to choose the distance measure that best classfied the pre-defined categories.[13] The ROC tests also show that one of the four clusters classifies GWF monarchies with very good precision (AUC=0.95), while the clusters closest to military (0.81) and party (0.83) regimes classify fairly well. This indicates the clusters we identify match observations that GWF categorize as monarchies almost exactly and military and party regimes closely.

The cluster that classifies GWF personalist regimes the best does so with an AUC of 0.68. This suggests that there is not as much overlap between this cluster and GWF personalist regimes, as there is between the other clusters and the respective GWF categories. When we inspect the cluster data for regimes identified by GWF as personalist, we find that many of them fall into the military (or party) cluster in the first years of the regime and then move to the personalist cluster, reflecting

[12]Figure A-7 in the appendix shows how the observations that fall into each cluster map onto the latent dimensions identified in the dimension analyisis.

[13]We tested the following distance measures: absolute value, correlation similarity, Euclidean distance (squared), matching, Jaccard, and Gower. We chose clusters from the Gower measure because they performed best in the ROC tests. We also tested three hierachical clustering algorithms (average, centroid, and complete) but none outperformed k-medians clusters in the ROC tests. The k-medians approach also performed better than the k-means clustering.

the consolidation of a leader's power.[14] That many dictatorships previously identified by GWF as personalist start out as military or party regimes, according to the cluster analysis, explains why the observations in the GWF personalist category in the upper right plot in Figure 6 are observed with both high and low values of dimension 3 (personalism) instead of being grouped together above zero on the vertical axis (i.e. high level of personalism). The latent dimensions approach to measuring regimes and the categories from the cluster analysis both employ time-varying data and thus pinpoint changes in the level of personalism over time within a particular regime.

Latent Dimensions in a Model of Regime Breakdown

In this section, we present an empirical model of regime breakdown similar to one in Geddes (2003), using updated data on autocratic regimes from Geddes, Wright and Frantz (2014). First, we test a model of regime failure with the binary categories of regime type (personalist is the omitted category) as explanatory variables; then we substitute three cluster variables derived from the raw data on regime characteristics: cluster 2 (monarchy); cluster 3 (party); and cluster 4 (military), with cluster 1 (personalist) omitted. The third model replaces the GWF regime categories with the three strongest latent dimensions from the dimension analysis (dimensions 1, 2, and 3).

Results from the first model are consistent with main findings in Geddes (1999) and Geddes (2003): military regimes are less stable than personalist ones, while party regimes and monarchies are more stable. The next column shows that regimes in *Cluster 4*, which is the cluster we believe most closely matches *Military regime*, are less stable than regimes in *Cluster 1* (omitted). The estimates for the coefficents for *Cluster 3* and *Cluster 2* – which we believe correspond to *Party regimes* and *Monarchies*, respectively – are both negative but only statistically significant at the 0.08 level.[15] Nonetheless, these estimates show that cluster variables yield similar results to those obtained when employing the GWF regime type variables.

[14] We find this, for example, in: Kerekou's regime in Benin, Conte's regime in Guinea, Traore's regime in Mali, the Kim regime in North Korea, Eyadema's regime in Togo, and Saleh's regime in Yemen.

[15] The standardized coefficient estimates from a linear probability model are: *GWF party* (-0.076); *GWF military* (0.118); *GWF monarchy* (-0.063); *Cluster 1* (-0.024); *Cluster 2* (0.125); *Cluster 4* (-0.031); *Dimension 1* (-0.097); *Dimension 2* (0.065); and *Dimension 3* (-0.063).

Table 3: Autocratic regime breakdown

	(1)	(2)	(3)
GWF party regime	-0.379*		
	(0.09)		
GWF military regime	0.489*		
	(0.10)		
GWF monarchy	-0.519*		
	(0.16)		
Cluster 1 (party)		-0.162	
		(0.09)	
Cluster 2 (military)		0.535*	
		(0.11)	
Cluster 4 (monarchy)		-0.234	
		(0.14)	
Dimension 1 (party)			-0.178*
			(0.04)
Dimension 2 (military)			0.128*
			(0.04)
Dimension 3 (personalism)			-0.099*
			(0.04)
Log GDP pc_{t-1}	-0.055	-0.045	-0.061
	(0.04)	(0.04)	(0.04)
Economic $\text{growth}_{t-1,t-2}$	-1.368	-1.277	-1.196
	(0.70)	(0.68)	(0.66)
Civil war_{t-1}	0.066	0.100	0.096
	(0.08)	(0.07)	(0.07)
(Intercept)	-1.256*	-1.384*	-1.249*
	(0.34)	(0.33)	(0.34)
Area under ROC	0.722	0.696	0.688

* p<0.05. 259 regimes in 114; 4169 observations. Years:1949-2010. Omitted category in column (1) is *Personalist regime*; in column (2), Regime 1 (personalist). All columns include the following control variables (not reported): decade fixed effects and regime duration cubic polynomials. Probit with standard errors clustered on regime in parentheses.

The final column shows that the estimates from the latent dimensions are also consistent with those reported in the first column: estimates for *Dimension 1* and *Dimension 3*, which we believe measure party strength and personalism, are negative and significant, indicating these features of autocracies stablize them. The estimate for *Dimension 2*, which we believe measures military strength, indicates that this characteristic is associated with a higher likelihood of regime failure. These models illustrate that the clusters and latent dimensions derived from the raw data perform as expected in a standard empirical application.

In Table 4, we compare how the first dimension, which we believe measures party strength,

Table 4: Comparing *Inclusive* and *dimension 1*

	(4)	(5)	(6)
Inclusive	-0.022		
	(0.04)		
Dimension 1		-0.210*	
		(0.04)	
Linear prediction			-0.054
			(0.07)
Residuals			-0.292*
			(0.05)
Dimension 2	0.119*	0.106*	0.117*
	(0.04)	(0.04)	(0.04)
Dimension 3	-0.138*	-0.104*	-0.124*
	(0.05)	(0.04)	(0.04)
GDP pc (log)	-0.080	-0.103*	-0.120*
	(0.05)	(0.05)	(0.05)
Growth	-1.028	-0.922	-0.870
	(0.71)	(0.71)	(0.71)
Prior democracy	0.260*	0.245*	0.239*
	(0.10)	(0.09)	(0.09)
Civil war	0.071	0.074	0.072
	(0.07)	(0.08)	(0.07)
(Intercept)	-0.629	-0.607	-0.447
	(0.44)	(0.41)	(0.42)
Area under ROC	0.682	0.709	0.718

* $p<0.05$. 170 regime breakdown events in 244 regimes in 113 countries, 1951-1988 and 1991-2000 (3439 observations). All columns include the following control variables (not reported): decade fixed effects and regime duration cubic polynomials. Standard errors clustered on regime in parentheses.

performs relative to Coppedge's inclusiveness variable in the model of autocratic breakdown. First, we substitute *Inclusive* in model (3); this reduces the sample size because of missing values on this variable. The estimate for *Inclusive* is negative (the expected direction), but not statistically different from zero. Second, we test the model with *Dimension 1* but on the same smaller sample as the model with *Inclusive*. The estimate of *Dimension 1* is in the expected direction and statistically different from zero. Comparing coefficient sizes in a non-linear model is not usually appropriate but we note that the standardized coefficient of interest in model (4) (-0.025) is smaller than the respective estimate in (5) (-0.210).

Last, we regress *Dimension 1* on *Inclusive* to create two measures, one from the predicted values (*Linear prediction*, which is perfectly correlated with *Inclusive*) and one from the residuals

(*Residuals*, which is uncorrelated with *Inclusive*). We then add both to the model instead of *Inclusive* or *Dimension 1*. In model (6), *Linear prediction* is small and not statstically different from zero but *Residuals* is in the expected direction and statistically significant. This indicates the variation in *Dimension 1* that is uncorrelated with *Inclusive* accounts for the expected negative correlation between *Dimension 1* and autocratic breakdown. Thus, although the first dimension is correlated with a measure of inclusiveness, this co-variation does not contain the information that correlates with regime breakdown in the expected direction.

Discussion

In this paper, we introduce the problems researchers face in attempting to measure important features of autocratic rule, and then construct an initial measure of three latent dimensions of autocratic rule that contain substantially different information than existing measures of democracy. We show that the first dimension captures the strength of party rule while the second measures the military's institutional strength vis-a-vis the party and the leader, while a third dimension measures personalist power. This approach, we argue, improves upon existing categorical typologies by producing continuous, time-varying measures of latent concepts.

We plan to re-construct the latent dimensions of autocratic rule when the raw data collection is complete. Further, we can extend this research by incoporating information from other data sets that either measure additional features of autocratic rule – such as the level of institutionalization (e.g. Gandhi (2008)), the degree of personalism (e.g. Hadenius and Teorell (2007)) and the type of military involvment (e.g. Svolik (2012)) – or that identify alternative typologies (Hadenius and Teorell 2007; Cheibub, Gandhi and Vreeland 2010). This step is crucial for specifying the structure of autocratic rule because we cannot currently make any formal claims beyond bivariate correlations about whether the information in these other data sets is already contained in the detailed historical data used in this project or if this additional information will identify new dimensions important to autocratic rule.

Finally, our initial attempt to structure the data uses principal components analysis, which relies on a linear projection of the many-dimensional features into the lower-dimensional ones.

More flexible approaches – such as self-organizing maps – may be fruitful as well, especially when there is structural missingness or non-linearity in the scales for some variables in the data.

References

Aksoy, Deniz, David B Carter and Joseph Wright. 2012. "Terrorism in Dictatorships." *The Journal of Politics* 74(3):810–826.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh. 2001. "New tools in comparative political economy: The Database of Political Institutions." *The World Bank Economic Review* 15(1):165–176.

Bratton, Michael and Nicolas van de Walle. 1997. *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective*. New York: Cambridge University Press.

Brooker, Paul. 1995. *Twentieth-Century Dictatorships: The Ideological One-Party States*. New York: New York University Press.

Brooker, Paul. 1997. *Defiant Dictatorships: Communist and Middle Eastern Dictatorships in a Democratic Age*. Basingstoke, Hampshire: Macmillan.

Brooker, Paul. 2000. *Non-Democratic Regimes: Theory, Government and Politics*. New York: St. Martin's Press.

Brownlee, Jason. 2009. "Portents of pluralism: how hybrid regimes affect democratic transitions." *American Journal of Political Science* 53(3):515–532.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson and James Morrow. 2003. *Logic of Political Survival*. Cambridge: MIT Press.

Bunce, Valerie and Sharon Wolchik. 2011. *Defeating Authoritarian Leaders in Post-Communist Countries*. New York: Cambridge University Press.

Chehabi, H. E. and Juan J. Linz. 1998. *Sultanistic Regimes*. Baltimore, MD: The Johns Hopkins University Press.

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

Clapham, Christopher. 1988. *Transformation and Continuity in Revolutionary Ethiopia*. Cambridge: Cambridge University Press.

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.

Coppedge, Michael, Angel Alvarez and Claudia Maldonado. 2008. "Two persistent dimensions of democracy: Contestation and inclusiveness." *Journal of Politics* 70(3):632–647.

Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I Lindberg, Kelly McMann, Pamela Paxton et al. 2011. "Conceptualizing and measuring democracy: A new approach." *Perspectives on Politics* 9(02):247–267.

Diamond, Larry Jay. 2002. "Thinking About Hybrid Regimes." *Journal of democracy* 13(2):21–35.

Erlich, Haggai. 1983. "The Ethiopian Army and the 1974 Revolution." *Armed Forces and Society* 9(3):455–81.

Fjelde, Hanne. 2010. "Generals, Dictators, and Kings Authoritarian Regimes and Civil Conflict, 1973 – 2004." *Conflict Management and Peace Science* 27(3):195–218.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship.* New York: Cambr. Univ. Press.

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(1):forthcoming.

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.

Gurses, Mehmet and T. David Mason. 2010. "Weak states, regime types, and civil war." *Civil Wars* 12(1-2):140–155.

Hadenius, Axel and Jan Teorell. 2007. "Pathways from authoritarianism." *Journal of Democracy* 18(1):143–157.

Haile-Selassie, Teferra. 1997. *The Ethiopian Revolution, 1974-1991: From a Monarchical Autocracy to a Military Oligarchy.* London: Kegan Paul International.

Honaker, James and Gary King. 2010. "What to do About Missing Values in Time Series Cross-Section Data." *American Journal of Political Science* 54(2):561–581. http://gking.harvard.edu/files/abs/pr-abs.shtml.

Honaker, James, Gary King and Matthew Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45(7). http://www.jstatsoft.org/v45/i07/paper.

Howard, Marc Morjé and Philip G Roessler. 2006. "Liberalizing electoral outcomes in competitive authoritarian regimes." *American Journal of Political Science* 50(2):365–381.

Huntington, Samuel P. 1968. *Political Order in Changing Societies.* New Haven: Yale University Press.

Huntington, Samuel P. 1991. "How Countries Democratize." *Political Science Quarterly* 106(4):579 – 616.

Israel, John. 1974. The Red Guards in Historical Perspective: Continuity and Change in the Chinese Youth Movement. In *Communist Systems in Comparative Perspective*, ed. Lenard J. Cohen and Jane P. Shapiro. Garden City, NY: Anchor Press/Doubleday.

Janowitz, Morris. 1960. *The Professional Soldier : A Social and Political Portrait.* New York: Free Press.

Karl, Terry Lynn. 1995. "The Hybrid Regimes of Central America." *Journal of Democracy* 6(3):72–86.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69. http://gking.harvard.edu/files/abs/evil-abs.shtml.

Klein, Donald W. and Lois B. Hager. 1974. The Ninth Central Committee. In *Communist Systems in Comparative Perspective*, ed. Lenard J. Cohen and Jane P. Shapiro. Garden City, NY: Anchor Press/Doubleday.

Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War.* Cambridge University Press.

Linz, Juan J. 2000. *Totalitarian and Authoritarian Regimes.* Boulder: Lynne Reinner Publishers.

Linzer, Drew A. and Jeffrey K. Staton. 2011. A Measurement Model for Synthesizing Multiple Comparative Indicators: The Case of Judicial Independence. In *presentation at the 2011 Annual Meeting of the American Political Science Association, September.* pp. 1–4.

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.

Nordlinger, Eric. 1977. *Soldiers in politics: military coups and governments.* Englewood Cliffs, NJ: Prentice Hall.

Peceny, Mark, Caroline C. Beer and Shannon Sanchez-Terry. 2002. "Dictatorial Peace?" *American Political Science Review* 96(1):15–26.

Pemstein, Daniel, Stephen A.M̃eserve and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4):426–449.

Ristaino, Marcia R. 1987. Chapter 10 – Party and Government. In *China: A Country Study*, ed. Robert L. Worden, Andrea Matles Savada and Ronald Dolan. Washington, DC: Federal Research Division, Library of Congress. `http://lcweb2.loc.gov/frd/cs/cntoc.html` [Accessed: 11 August 2014].

Robinson, Thomas W. 1972. Lin Piao as an Elite Type. In *Elites in the People's Republic of China*, ed. Robert A. Scalapino. Seattle, WA: University of Washington Press.

Rubin, Donald. 1976. "Inference and Missing Data." *Biometrika* 63:581–592.

Scalapino, Robert A. 1972. The Transition in Chinese Party Leadership: A Comparison of the Eighth and Ninth Central Committees. In *Elites in the People's Republic of China*, ed. Robert A. Scalapino. Seattle, WA: University of Washington Press.

Schafer, Joseph L. 1997. *Analysis of incomplete multivariate data.* London: Chapman & Hall.

Schedler, Andreas. 2009. "Electoral authoritarianism." *The SAGE Handbook of Comparative Politics. London* pp. 381–394.

Shinn, Rinn-Sup and Robert L.Worden. 1987. Chapter 1– Historical Setting. In *China: A Country Study*, ed. Robert L. Worden, Andrea Matles Savada and Ronald Dolan. Washington, DC: Federal Research Division, Library of Congress. `http://lcweb2.loc.gov/frd/cs/cntoc.html` [Accessed: 11 August 2014].

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

Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–217.

Weber, Max. 1964. *The Theory of Social and Economic Organization.* New York: Free Press.

Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1):35–64.

Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106(2):326–347.

Wilson, Matthew C. 2014. "A Discreet Critique of Discrete Regime Type Data." *Comparative Political Studies* 47(4):forthcoming.

Wilson, Matthew C. and James A. Piazza. 2014. "Autocracies and Terrorism: Conditioning Effects of Authoritarian Regime Type on Terrorist Attacks." *American Journal of Political Science* Forthcoming.

Wintrobe, Ronald. 1990. "The Tinpot and the Totalitarian: An Economic Theory of Dictatorship." *American Political Science Review* 83(4):849 – 872.

Wright, Joseph. 2008. "Do Authoritarian Institutions Constrain? How Legislatures Affect Economic Growth and Investment." *American Journal of Political Science* 52(2):322–343.

Wright, Joseph. 2009. "How foreign aid can foster democratization in authoritarian regimes." *American Journal of Political Science* 53(3):552–571.

Appendix

Raw data variable definitions

These definitions are not the same as the coding rules, which are much more detailed. The defintions are only intended to provide readers with a basic understanding of the type of information contained in the raw data.

- `leaderrole` (treated as categorical)
  identify how the regime leader achieved office and/or whose support put him in office
  - prior democracy
  - dominant party
  - military junta/executive committee
  - insurgency
  - traditional hereditary succession
  - (clerical, civilian, interim, foreign)
- `seizure` (treated as categorical)
  how did the regime obtain power
  - hereditary succession
  - military coup
  - insurgency/rebels
  - foreign (omitted)
  - uprising
  - election
  - authoritarian succession
- `supportparty` (binary)
  0. no support party
  1. support party
- `partyleader` (ordinal)
  0. no support party or party leader
  1. party leader is regime leader
  2. party leader is relative of regime leader
  3. party leader selected by regime leader
  4. party leader selection influenced by regime leader
  5. party leader is not regime leader and selection controlled by group that excludes the regime leader
- `partyhistory` (ordinal)

0. no support party
1. party created after seizure of power
2. prior party, created to support autocratic leader election (e.g. Fujimori)
3. prior party, never won electoral support
4. prior party, won support under prior autocracy
5. insurgent/rebel party
6. 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. cabinent ministers (except defense) are party members

- `partymilit` (ordinal)
  missing if no party or military

  0. military controls 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` (ordinal)

  0. not missing on `partymilit`
  1. no party or military (missing on `partymilit`)

- `partyexcom` (ordinal)

  0. no support party
  1. regime leader chooses party executive committee
  2. faction that supports the regime leader dominates the party executive committee
  3. 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` (ordinal)

  0. no support party
  1. party leadership is dominated by people from particular ethnic/region/religious
  2. party leadership is multi-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

- `cabciv` (ordinal)

  0. most important cabinet positions held by military
  1. cabinet is civilians or insurgents, but some military in positions other than defense
  2. civilian cabinet (except defense)

- `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

- `milconsult` (binary)

  0. no consulative 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

- `milmerit` (ordinal)

  0. regime leader promotes officers loyal to himself or from his ethnic, tribal, regional, or religious group; and leader does forces officers not from his group into retirement for political reasons; or no/foreign military
  1. some evidence of promotions of top officers based loyalty to the regime leader or from his group, but no reports of widespread use of loyalty as a criterion from promotion or retirement
  2. regime leader does *not* promote officers loyal to himself or from his ethnic, tribal, regional, or religious group; and leader does *not* force officers not from his group into retirement for political reasons

- `militparty` (treated as categorical)
  identify how the regime leader achieved office and/or whose support put him in office

  – prior democracy
  – dominant party
  – military junta/executive committee
  – insurgency
  – traditional hereditary succession

- `milethnic` (treated as categorical)
  identify how the regime leader achieved office and/or whose support put him in office

  - regime leader's supporters are not organized in a party (no party)
  - regime leader allies with a pre-existing party (ally party)
  - regime leader or a close ally creates a party to support the regime after his accession to office (newp arty)
  - regime leader's supporters were organized into a party prior to his accession to power, and that party now supports the regime (prior party)
  - (regime leader is not in military)

- `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)

  - chosen by family (omitted)
  - elected in prior dictatorship
  - elected in one candidate election
  - elected in one party election
  - selected by legislature elected in multiparty elections
  - elected in multiparty elections
  - elected in prior democracy

- `legnoms` (ordinal)

  0. no opposition in legislature; or no legislative body
  1. legislature selected by indirect election from lower body; comprised of local/tribal notables; or selected by regime insiders as societal representatives
  2. opposition allowed to contest but regime holds veto power of candidate selection; and regime leader or committee has veto power over ruling party candidate selections
  3. opposition allowed to contest but regime holds veto power of candidate selection; and support party candidate selection influenced by local party leaders or regime factions
  4. opposition or independents allowed to contest; and ruling party candidate selection influenced by local party leaders or faction members
  5. legislature chosen in prior democratic regime or competitive pre-independence election

- `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

- `paramil` (ordinal)

  0. no paramilitary; no military; or officers are mostly foreigners
  1. paramilitary forces created to fight civil war on regime's side
  2. party militia or paramilitary organized by dominant party
  3. regime leader creates paramilitary forces, a president's guard, or new security forces apparently loyal to himself

- `secty_app` (treated as categorical)

  – security apparatus controlled by military (omitted)
  – security apparatus controlled by dominant party
  – security apparatus controlled personally by regime leader

Summary Statistics

Table A-1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
supportparty	0.726	0.446	0	1	4549
ldr_priordem	0.081	0.273	0	1	4549
ldr_domparty	0.262	0.44	0	1	4549
ldr_military	0.288	0.453	0	1	4549
ldr_insurgency	0.099	0.299	0	1	4549
ldr_hereditary	0.112	0.315	0	1	4549
partyhistory	2.889	2.513	0	6	4549
partymins	1.794	1.335	0	3	4549
partyexcom	1.126	0.885	0	3	4549
partymilit	1.064	1.382	0	4	4549
partymilit2	0.545	0.45	0	1	4549
localorgzns	1.365	0.883	0	2	4549
excomcivn	1.435	1.252	0	3	4549
multiethnic	1.239	0.854	0	2	4549
heirparty	0.379	0.445	0	1	4549
heirfamily	0.423	0.44	0	1	4549
legcompetn	4.278	2.839	0	8	4549
leaderciv	0.563	0.462	0	1	4549
heirciv	1.07	0.867	0	2	4549
cabciv	1.319	0.615	0	2	4549
militrank	0.998	1.553	0	4	4549
milconsult	0.125	0.278	0	1	4549
milmerit	0.703	0.646	0	2	4549
militparty_newparty	0.138	0.302	0	1	4549
militparty_allyparty	0.051	0.173	0	1	4549
militparty_noparty	0.134	0.302	0	1	4549
militparty_priorparty	0.02	0.101	0	1	4549
milethnic_dom	0.454	0.406	0	1	4549
milethnic_hetero	0.403	0.391	0	1	4549
milethnic_homo	0.144	0.264	0	1	4549
ldrrotation	0.055	0.182	0	1	4549
partyrbrstmp	0.321	0.403	0	2	4549
heirclan	0.32	0.387	0	1	4549
officepers	0.671	0.378	0	1	4549
paramil	1.471	1.071	0	3	4549
sectyapp_party	0.136	0.241	0	1	4549
sectyapp_pers	0.66	0.371	0	1	4549
electldr_notelect	0.354	0.438	0	1	4549
electldr_priordict	0.013	0.084	0	1	4549
electldr_1candidate	0.192	0.352	0	1	4549
electldr_1party	0.072	0.216	0	1	4549
electldr_multileg	0.08	0.232	0	1	4549
electldr_multiexec	0.198	0.356	0	1	4549
electldr_priordem	0.038	0.16	0	1	4549
legnoms_indirect	0.089	0.2	0	1	4549
legnoms_veto	0.281	0.361	0	1	4549
legnoms_noveto	0.23	0.331	0	1	4549
legnoms_priordem	0.045	0.135	0	1	4549
seizure_coup	0.303	0.46	0	1	4549
seizure_rebel	0.248	0.432	0	1	4549
seizure_family	0.049	0.215	0	1	4549
seizure_election	0.135	0.342	0	1	4549

Table A-1 gives summary statistics of the observed data for the variables described above. For the dimension and cluster analysis, we reduce the number of variables to the following: support-

party, ldr_domparty, ldr_military, ldr_hereditary, partyhistory, partymins, partyexcom, partymilit, partymilit2, localorgzns, excomciv, multiethnic, heirparty, leaderciv, heirciv, milconsult, militrank, militparty_newparty, militparty_allyparty, militparty_noparty, partyrbrstmp, officepers, paramil, sectyapp_party, sectyapp_pers, electldr_notelect, seizure_coup, seizure_family, electldr_1candidate, legcomp. This reduction increases the explained variation in the first 3 dimensions from 53 percent to 81 percent, while still yielding dimensions that are highly collinear with the dimensions drawn from analysis of the full data set: 0.99 for dimension 1; 0.98 for dimension 2; and 0.94 for dimension 3.

Missing data imputations

The number of complete observations for each variable were given above in Table A-1. The patterns within the missing values are shown below in A-1. Variables are aligned left-to-right by the number of observed values.

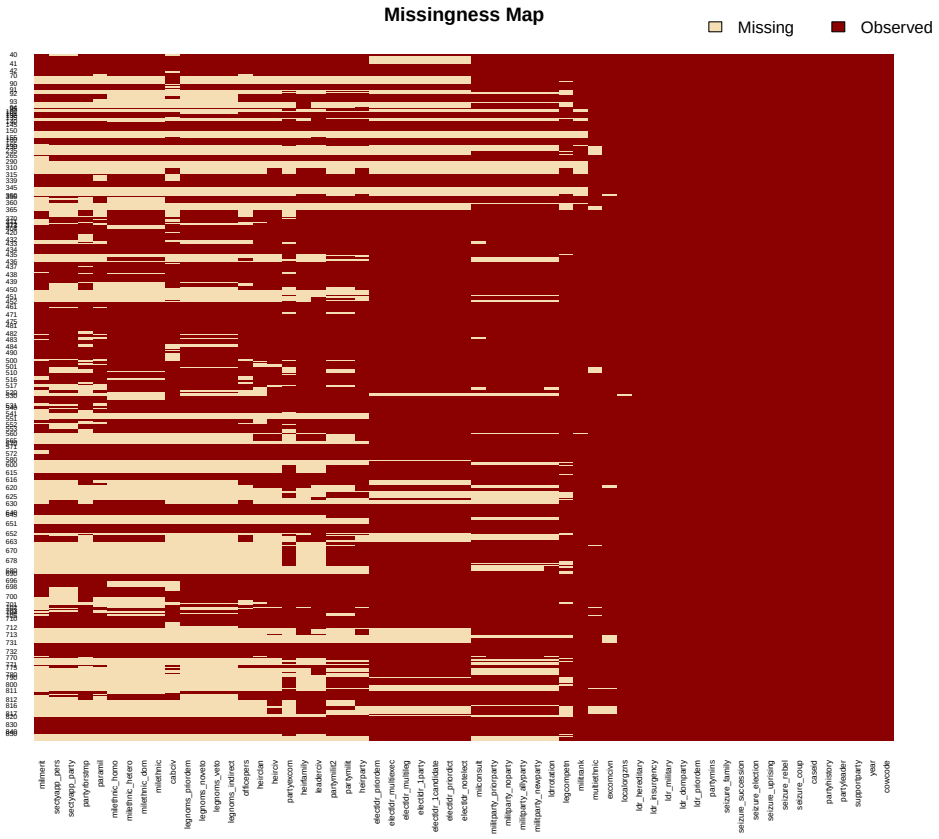


Figure A-1: Missingness patterns in autocratic indicators.

Principle components

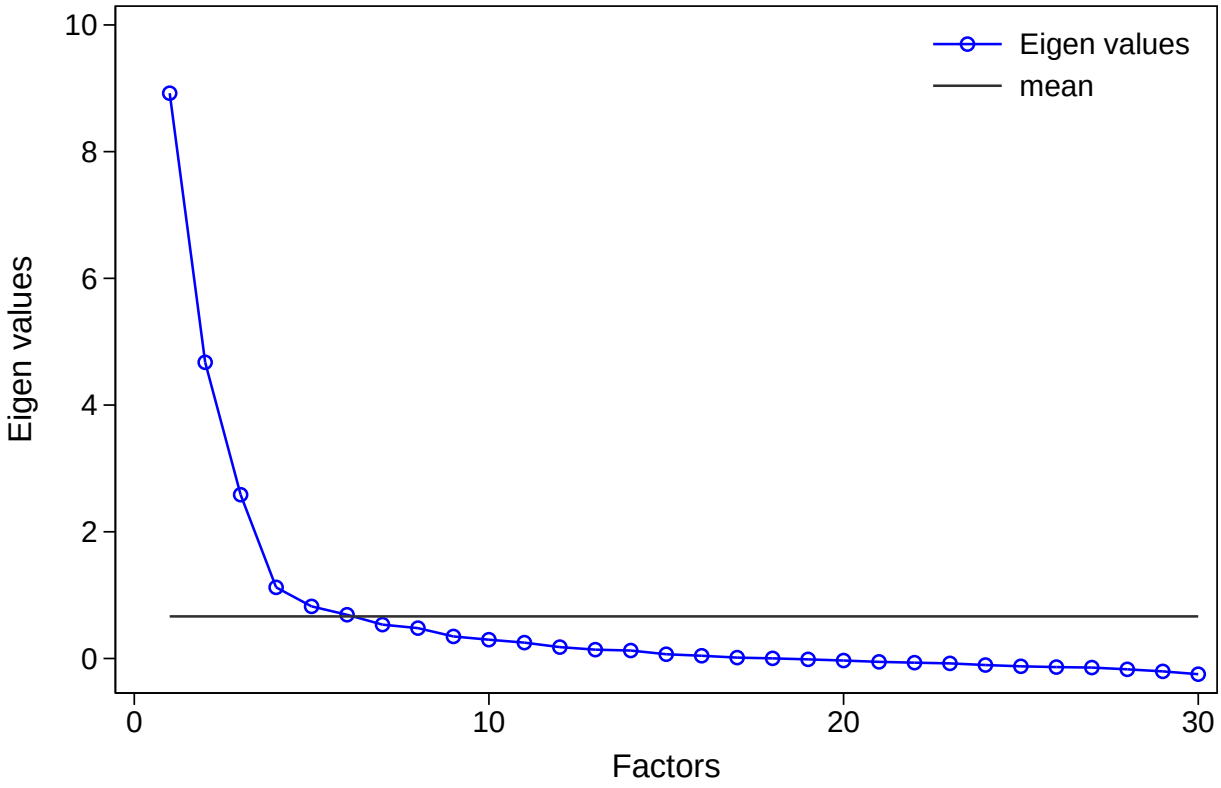


Figure A-2: Eigenvalues for 30 dimensions. The first 3 dimensions explain 81 percent of the variation in the 30 variables.

Table A-2: Rotation

Variables	dimension 1	dimension 2	dimension 3
supportparty	0.8292	0.3938	0.3146
ldr_domparty	0.5171	-0.0718	-0.1301
ldr_military	-0.5253	0.6712	-0.0924
ldr_hereditary	-0.3518	-0.6426	0.0656
partyhistory	0.7974	0.0802	-0.0417
partymins	0.8144	0.3087	0.1348
partyexcom	0.8	0.3029	-0.0004
partymilit	0.7053	0.0437	-0.3402
partymilit2	-0.7205	-0.2685	0.1875
localorgzns	0.8446	0.3574	0.2474
excomcivn	0.7429	0.1837	0.0204
multiethnic	0.7877	0.3346	0.237
heirparty	0.7521	-0.0004	-0.2244
leaderciv	0.3663	-0.7353	-0.0491
heirciv	0.4458	-0.6042	-0.1493
milconsult	-0.4394	0.493	-0.308
militrank	-0.4794	0.7159	-0.0066
militparty_newparty	-0.1299	0.5464	0.4596
militparty_allyparty	-0.0742	0.3841	-0.0411
militparty_noparty	-0.6328	0.3763	-0.4131
partyrbrstmp	0.6422	0.0326	-0.3366
officepers	-0.1727	-0.1626	0.5949
paramil	-0.0147	0.1768	0.446
sectyapp_party	0.334	0.0181	-0.4563
sectyapp_pers	-0.1393	-0.0667	0.592
electldr_notelect	-0.1499	0.3519	-0.4451
seizure_coup	-0.4688	0.6177	0.038
seizure_family	-0.2325	-0.4628	0.0928
electldr_1candidate	0.0933	0.214	0.3454
legcompetn	0.3209	-0.1788	0.2487

Comparison with democracy measures

In Figure A-5 we present simple summary statistics of our autocratic measures of the latent dimension, across the range of polity scores. We only have 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. In Figure A-4 we saw that all 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 0, 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.

However, the variances, shown as the dashed lines, are systematically related to the polity dimension. Except at the lower bound, when again the data is tightly clustered, the variance of the dimensions slowly decreases as polity increases. This suggests that as states become more democratic, states are becoming more alike in the latent autocratic structure, and this dimension has less meaningful information.

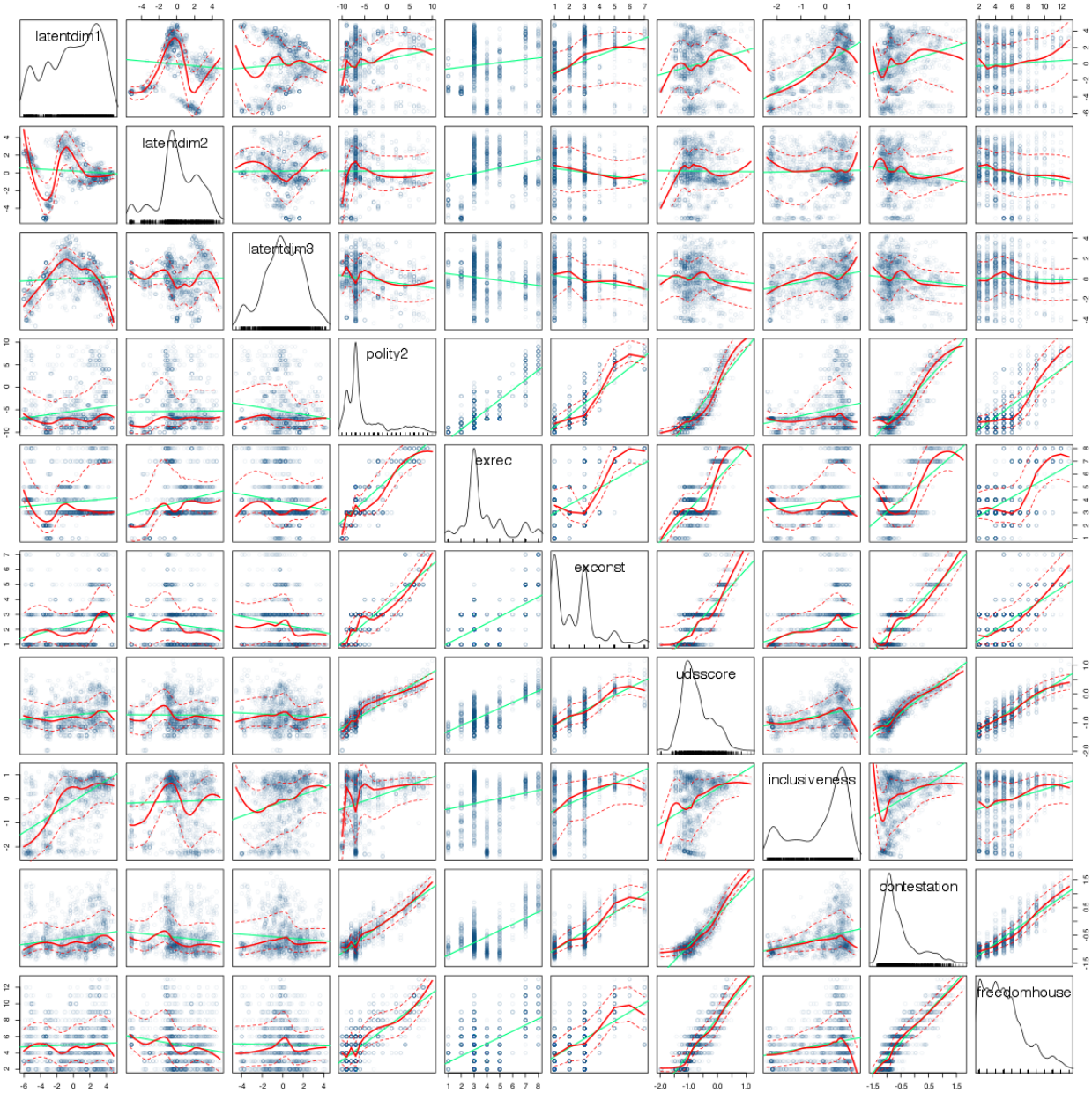


Figure A-3: Scatterplots of latent autocratic dimensions with alternate measures of democratization. The first three columns show our recovered dimensions (on the x-axis, against the democratization scores (on the y-axis) along with a regression line (green) and spline smoother (red))

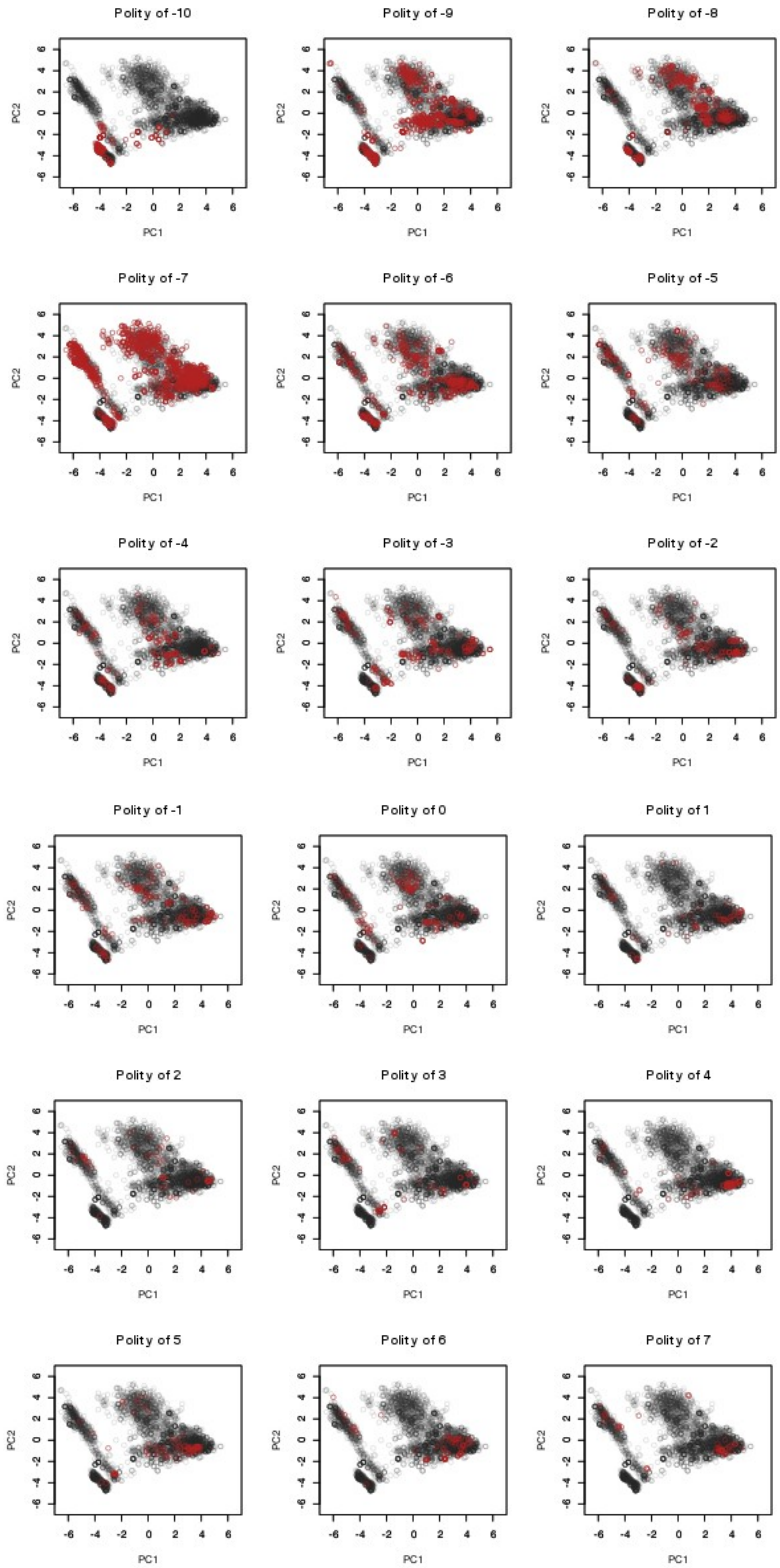


Figure A-4: A breakdown of our representation of the latent space, by Polity score. Gray points represent all data in the space to show the underlying distribution, while red points are the subset that have that particular level of polity score.

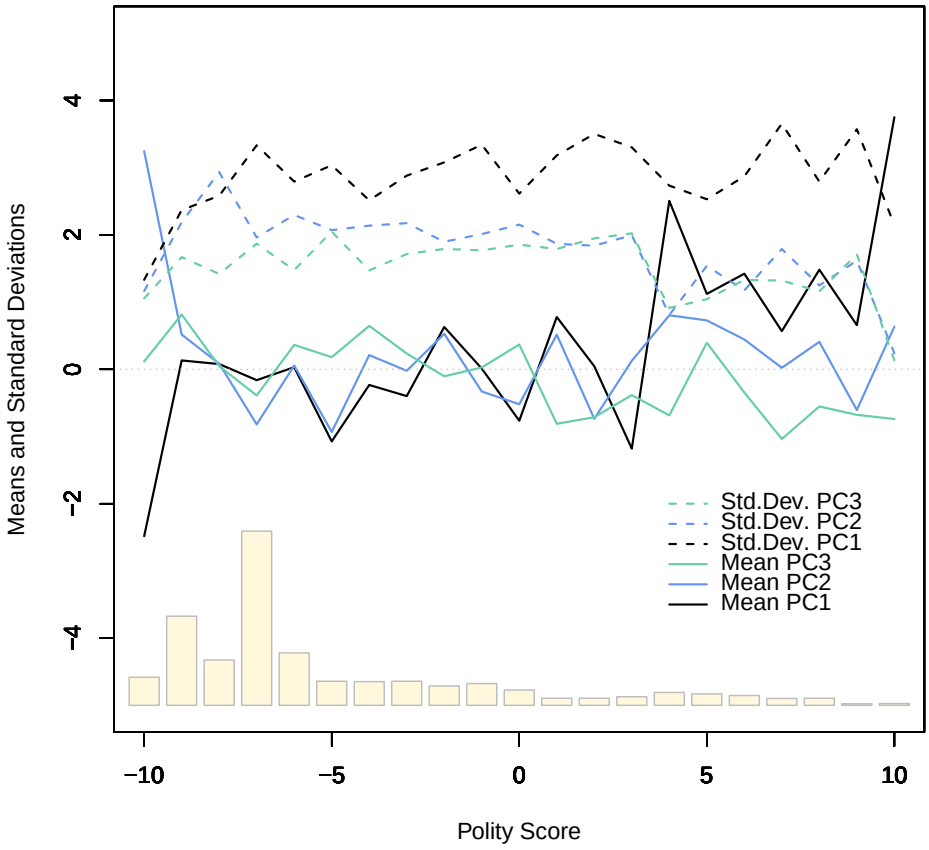


Figure A-5: Summary statistics of latent dimensions, across levels of democracy. Variances are decreasing across the range, while outside the very bounds of the data (-9 to 0), the mean positions appear unrelated to the Polity score.

Comparison with additional measures of autocracies

Table A-3: Pairwise bivariate correlations

Variables	dimension 1	dimension 2	dimension 3
Number of parties (Gandhi)	0.249	0.167	0.162
Number of parties (H&T)	0.099	0.134	0.043
Number of parties (DPI)	0.425	-0.021	0.232
Number of parties (Svolik)	0.287	0.214	0.149
One or more parties (Gandhi)	0.550	0.257	0.195
One or more parties (H&T)	0.609	0.059	0.242
One or more parties (DPI)	0.543	-0.040	0.249
One of more parties (Svolik)	0.560	0.322	0.161
Leader associated w. party (Svolik)	0.589	0.171	0.131
Institutions (Gandhi)	0.436	0.025	0.237
Legislative competitive index (Svolik)	0.339	-0.075	0.215
Legislative competitive index (DPI)	0.382	-0.052	0.251
Military regime (Gandhi)	-0.433	0.735	0.072
Military regime (H&T)	-0.471	0.599	-0.018
Military regime (DPI)	-0.430	0.710	0.165
Military regime (Svolik)	-0.481	0.703	0.040
Corporate military (Svolik)	-0.379	0.430	-0.045
Personal military (Svolik)	-0.227	0.447	0.104
Indirect military (Svolik)	-0.069	0.105	-0.035

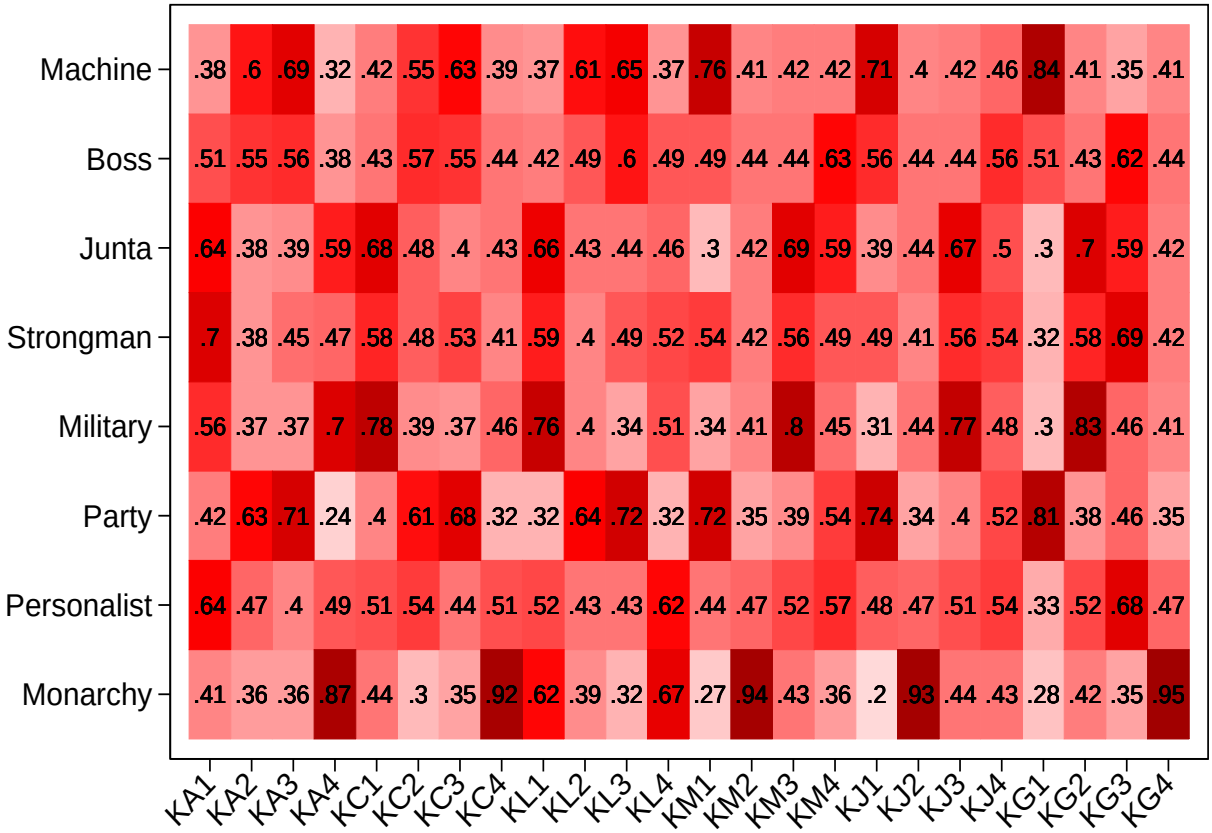


Figure A-6: ROC test scores using clusters to classify prior GWF and Weeks regime categories.

Cluster analysis

Each cell in Figure A-6 reports the area under the ROC (AUC) when using a particular distance measure to construct clusters. Darker shades of red indicate better classification (i.e. a higher AUC). The labels on the horizontal axis correspond to a set of four clusters derived from different distance measures. Each distance measure is identified by the first two letters of the label and the number corresponds to the number for each cluster. For example, *KG1*, *KG2*, *KG3*, and *KG4* denote the four clusters derived from the k-medians partion cluster algorithm that uses the Gower (KG) distance measure. Cluster numbers 1-4 are assigned randomly and should not be interpreted substantively. KA≡ absolute value distance; KC≡ correlation coefficient similarity measure; KL≡ Euclidean distance; KM≡ matching similarity coefficient; KJ≡ Jaccard binary similarity coefficient; and KG≡ Gower's dissimilarity coefficient.

Figure A-7 shows how the three dimensions from the dimension analysis map onto the cluster categories. Each circle represents a country-year observation in the data set. The light gray circles show all observations in a 2 dimensional space (dimension 1 and dimension 2 in the left two columns; dimension 2 and dimension 3 in the right two columns). The black circles in each plot show the observations that fall into a particular cluster category.

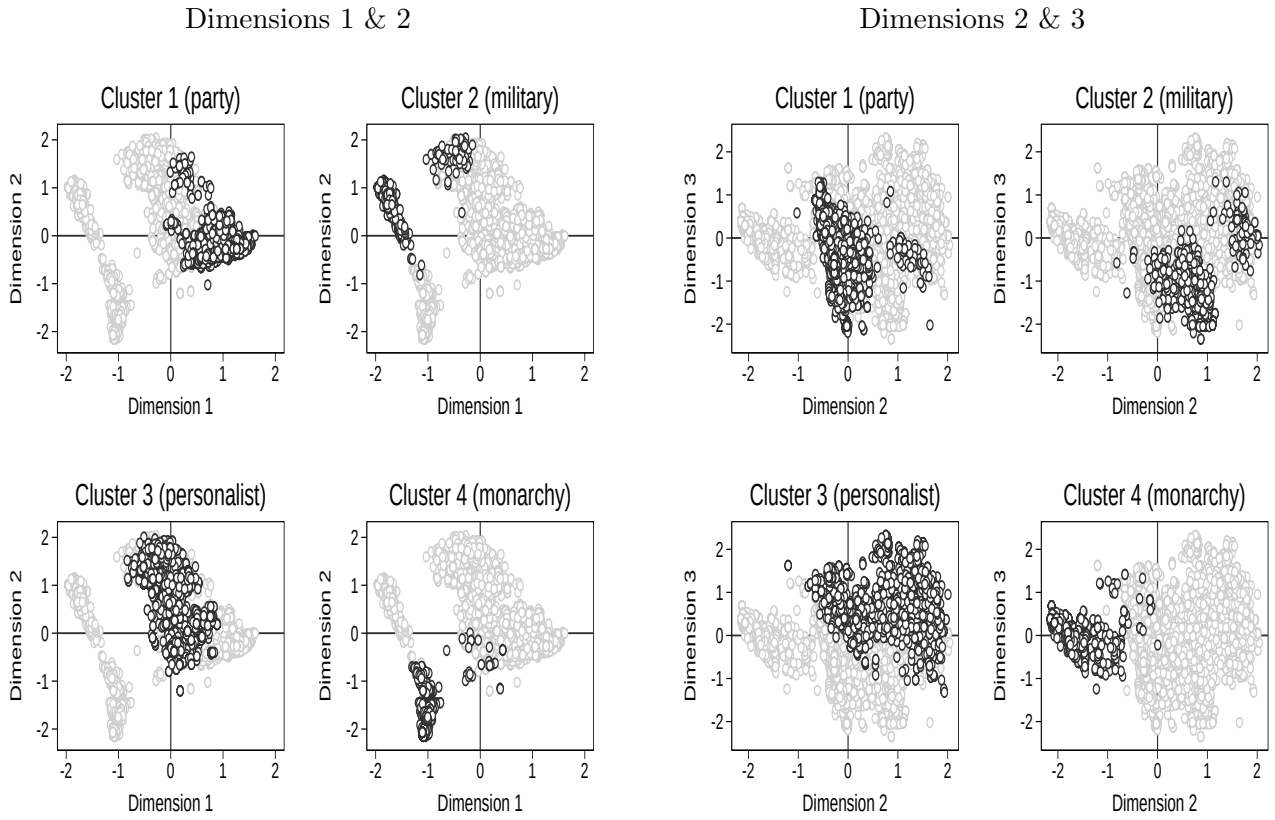


Figure A-7: Latent dimensions of autocratic rule, by clusters.