



university of
 groningen

faculty of behavioural
 and social sciences

Civil Liberties in Personalising Autocracies

How personalism degrades civil liberties by disempowering repressors

Author: Siebren Kooistra
Student Number: 3959627
Programme: BSc Sociology, Rijksuniversiteit Groningen
Supervisor: dr. Jacob Dijkstra
Second evaluator: dr. Rita Smaniotto

June 2022

Abstract:

This thesis investigates the way personalism (e.g., Geddes et al., 2018) affects four civil liberties. I theorise greater personalism to decrease freedom of expression through limiting perceived coup threat, decrease freedom of assembly because the dictator has less need for the general population to monitor the elite, and lower protection of life and physical integrity and freedom of movement because the competence of coercive institutions degrades. I test these hypotheses using multi-level (random intercept) models and find that more personalist countries have slightly lower civil liberties. The competence of coercive institutions operationalised as the rigour and impartiality of the public administration might mediate the relation between personalism both protection of life and physical integrity and freedom of movement, and possibly the relations to other civil liberties. My findings seem in line with Frantz et al. (2019), but substantively weak results and problems in accounting for dependence structures restrict my conclusions.

Table of Contents

Abstract:	2
Introduction:	5
Theory:	7
Freedom of expression: Transparency and shifting risks	7
Freedom of assembly: Who monitors the elite?	9
Protection of life and physical integrity and freedom of movement: The competence of repressors	10
Methods:	12
Dataset.....	12
Data collection procedure	12
Variable modelling.....	13
Lags and coding dates	13
Operationalisations of core variables.....	13
Personalism.....	14
Freedom of expression.....	15
Freedom of assembly	15
Protection of life and physical integrity.....	16
Freedom of movement	16
Rigour and impartiality public administration.....	16
Control variables and operationalisations.....	16
Regime type	17
International or internal conflict	17
Political violence.....	17
Population size	17
GDP per capita and GDP growth.....	18
Analysis plan.....	18

Results:.....	21
Univariate statistics.....	21
Bivariate statistics.....	24
Regression modelling.....	27
Freedom of expression.....	27
Freedom of assembly.....	30
Protection of life and physical integrity.....	33
Freedom of movement.....	35
Rigour and impartiality of the public administration.....	38
Conclusion & Discussion:.....	41
References:.....	44
Appendix: R script.....	50

Introduction:

The rule of one takes the effort of many and the compliance of more. This means that explaining such a rule of one, a personalised autocracy, requires a look at the interactions between a dictator, their elite, and the general populace. Previous work has already done much to explain how dictators gain personal power and consolidate it (e.g., Acemoglu et al., 2009; Geddes et al., 2018; Svobik, 2012). I will look into the consequences of personalism for four civil liberties, rights allotting freedom from various kinds of coercion or freedom to make certain choices (Davenport, 2007a).

Existing research into the civil liberties and personalism has looked into the effect of personalisation on the protection of life and physical integrity (Frantz et al., 2019; Greitens, 2016) and freedom of expression (Boleslavsky et al., 2021; Hollyer et al., 2019). I try to build on this knowledge, and will explicitly distinguish between freedom of expression, freedom of assembly, freedom of movement and protection of life and physical integrity. The processes creating personalised autocracy might affect different civil liberties in distinctive ways, making it theoretically meaningful to separate out the effects of personalism on various civil liberties (Davenport, 2007b; Møller & Skaaning, 2013a). To deepen knowledge on these effects, I will also test to what extent the relation between personalism and civil liberties is mediated by variables linked to the mechanisms I identify in my theory. My research question is:

“To what extent do the level of personalism and changes in the level of personalism within an autocratically ruled country affect respect for various civil liberties, and what are the mechanisms underlying these relations?”

To explain how personalised autocracy affects civil liberties, it is important to first establish what personalism is, and what perspective I will take on it. *Personalism* is the extent to which power in a regime is concentrated in the hands of a single individual (Geddes et al., 2018; Sinkkonen, 2021; Svobik, 2012). To explain how personalism evolves over time, I will use the coalition-formation model established by Acemoglu et al. (2008, 2009, 2012). This model starts from the assumption that autocratic regimes lack effective formal rules and procedures to regulate political actions (Gehlbach et al., 2016). This makes credible commitments hard, costly, or outright impossible (Acemoglu et al., 2008, 2009; Gehlbach et al., 2016; Svobik, 2012). When there is no trust in an independent law enforcement system (judiciary, police) for example, members of the elite with influence over parts of it will be

reluctant to give up this influence for fear of later being persecuted by that same system. Even if the person who receives power from others is benevolent, the fact that the power of this individual makes them impervious to external control makes others mistrustful. As such, those with access to power in an autocratic regime (the elite) feel like they must use it to have their interests met, and strive to increase their power for further security.

Because members of an autocratic elite strive to increase their power, they want a ruling coalition including themselves to be as small possible (Acemoglu et al., 2008). This means that members of a ruling coalition have an incentive to exclude others when they can safely do so. The incentive of members of the ruling coalition to shrink the ruling coalition creates the risk of factionalist infighting, and dictators tend to exploit factionalism in autocratic regimes by acting as an arbiter between factions (Geddes et al., 2018). By strategically choosing sides in internal conflicts, successful dictators eliminate powerful factions within the regime until the remaining factions are no longer able to limit this dictator (Geddes et al., 2018; Svobik, 2012). A dictator that has accrued some personal power can further reinforce factionalism by encouraging competition among government institutions and limiting the elite's opportunities to communicate and establish mutual trust (Geddes et al., 2018; Greitens, 2016). Still, leveraging internal conflict to personalise the regime fails more often than not, since (the rest of) the elite will eliminate a dictator if they comprehend the dictator's strategy on time (Svobik, 2012), while cohesive ruling coalitions lack the necessary factionalism (Geddes et al., 2018; see also Levitsky & Way, 2016). But people do not always have the foresight to avoid being outplayed (Jandoc & Juarez, 2019), and exogenous events or failed resistance can tip the balance of power (Acemoglu et al., 2008; Chin, 2020).

To summarise, previous research suggests that personalism results from the disruptions to the balance of power within an autocratic elite, and reduces protection of life and physical integrity and freedom of expression. I will broaden this research by considering not only protection of life and physical integrity and freedom of expression, but also freedom of assembly and freedom of movement. I will also try to deepen the research by looking at the mechanisms underlying relations between personalism and civil liberties.

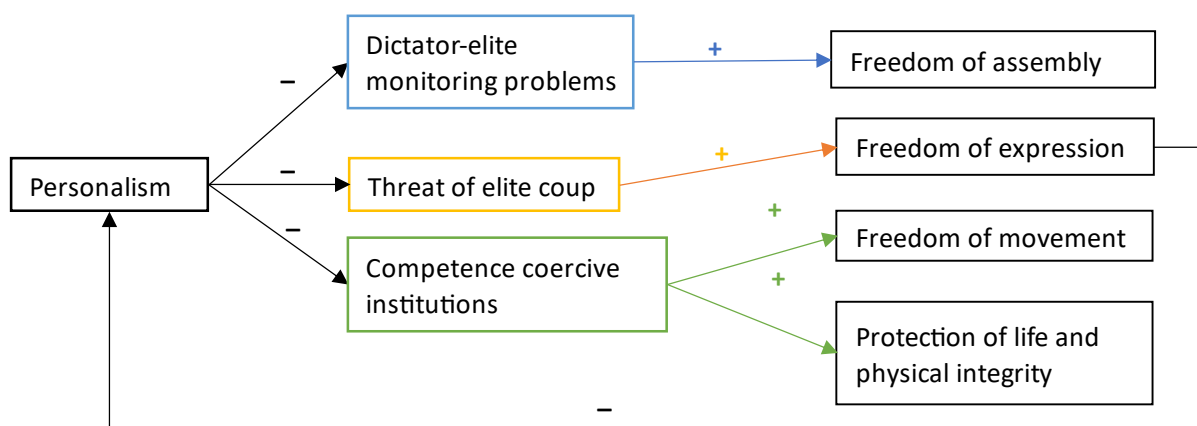
In the next section, I will detail the mechanisms through which personalism affects civil liberties. After this theoretical argument, I use a third section to introduce the dataset and my empirical approach. The fourth section gives descriptive statistics of the dataset and model results. I draw my conclusions and reflect on them in a fifth section, which rounds off the article.

Theory:

In this section, I will outline my hypotheses and the reasoning behind them. In doing this, I build on the general perspective presented in the introduction. My hypotheses are summarised in Figure 1.

Figure 1

Conceptual model of the relations between personalism and freedom of assembly, freedom of expression, freedom of movement and protection of life and physical integrity.



Note: Although omitted in the figure, it can be assumed that there are also direct arrows from personalism to civil liberties to account for other mechanisms linking personalism to civil liberties, even if these are not of interest to this study.

Freedom of expression: Transparency and shifting risks

Freedom of expression is the freedom to express, seek and receive thoughts and opinions in private or in public (Coppedge et al., 2021a, p. 46, p. 307; Møller & Skaaning, 2013b, p. 1071; Universal Declaration of Human Rights, 1948, Art. 19).

For the mechanism linking of personalism to freedom of expression and personalism, I propose that weak dictators maintain freedom of expression to keep their rivals in check, but sufficiently powerful dictators no longer need to do this and want to remove these checks when acting against their rivals. By and large, dictators mostly have an opportunity to establish themselves when the elite is factionalised and prone to infighting (Geddes et al., 2018). For the situation where a dictator is still relatively weak, this creates an initial environment where members of the elite have to be more wary of being ousted in infighting than having the regime come down due to popular protest. This is the case because members of the elite need to overcome less of a coordination problem than the general populace to rise up against a dictator (Svolik, 2012; Geddes et al., 2018). To keep the threat of an internal

power grab in check, members of the elite use the fact that cooperation problems for the general populace decrease when they can coordinate on shared knowledge of infighting in and thus vulnerability of the ruling coalition (Hollyer et al., 2019). To enable the creation of such shared knowledge, the dictator and other members of the elite can facilitate the general populace (journalists in particular) in observing and reporting infighting in the elite. While this poses a risk to the elite (including the dictator), this is outweighed by the fact that it discourages infighting that might lead to their own exclusion (Boleslavsky et al., 2021; Hollyer et al., 2019).

The effort required to restrict freedom of expression makes it unlikely that the dictator, or any subgroup of the elite for that matter, will unilaterally restrict freedom of expression as long as they do not completely dominate the coalition. Comprehensively limiting freedom of expression requires a great deal of coordination. And when it happens, it is likely to signal both to the rest of the elite and to the general populace that that subgroup is an urgent threat to their interests of retaining, respectively, power and freedom of expression. By unilaterally trying to limit freedom of expression, members of the ruling coalition incur the risk of facing coordinated resistance from both the rest of the ruling coalition and the general populace. Unless at least one of these threats can be removed, limiting freedom of expression makes the perpetrator's position untenable.

A dictator in a factionalised ruling coalition that is still vulnerable to coordinated resistance from the rest of the elite has a use for freedom of expression and incurs a lot of risk by limiting it. However, once their power becomes great enough to make resistance from within the elite ineffectual (see Svoboda, 2012), they have an opportunity to limit freedom of expression. Since their domination of the elite means that they no longer need external control to keep their rivals in check, their priority lies with limiting freedom of expression to disrupt coordination in the general populace (Boleslavsky et al., 2021; Hollyer et al., 2019). When moving to limit freedom of expression the dictator will still face popular resistance, but since the dictator can use considerable government resources it is likely that they can overcome this counterforce.

A point of note here is that the mechanism I sketched above not only implies that freedom of expression is limited after personalism is at sufficiently high levels, but also that limiting freedom of expression stimulates greater personalism. Decreasing freedom of expression takes away the constraints on infighting that drove the elite to accept some freedom of expression in the first place, and as such lower freedom of expression can amplify personalisation by permitting the dictator more aggressive use of their power dominance.

Empirically, I expect that from a factionalised ruling coalition that allows some freedom of expression, increasing personalism leaves freedom of expression unchanged at first but leads to a decrease in freedom of expression once the dictator has consolidated their position. I also expect there to be a reverse relation of less freedom of expression leading to increasing personalism, but I do not posit a specific hypothesis about this. This makes my hypothesis on freedom of expression:

H₁: *Freedom of expression in a country has a curvilinear association with personalism, remaining fairly constant at low levels of personalism and decreasing at high levels of personalism.*

Freedom of assembly: Who monitors the elite?

Freedom of assembly is the freedom to peacefully gather to signal opinions or concerns in public (Coppedge et al., 2021a, p. 225; Møller & Skaaning, 2013b, p. 1071).

My theory on the relation between personalism and freedom of assembly is based on the ability of a dictator to monitor their elite. Once in power, dictators need to monitor their elites both to make sure that they carry out their policies and to prevent any members of the elite from illicitly building an independent power base. For monitoring and disciplining their officials, dictators can use freedom of assembly by justifying punishment of threatening officials with public discontent. In addition, giving the general populace some opportunities to voice their discontents grants the dictator a source of information about unusual or dysfunctional behaviour among the executive elite (Geddes, 2018; Zu, 2020). However, the downside of freedom of assembly is that protests can be targeted against the dictator and their government as a whole as well as against specific policies. The dictator will want to prevent such protests, and repress them when they do occur (Davenport, 2007a; Zu, 2020). A dictator thus faces a trade-off between leaving some freedom of assembly to monitor their elite, and further limiting it to minimise the risk of protest against the regime as a whole.

This trade-off makes the creation of parallel coercive institutions and overlap in roles very important (Greitens, 2016). It is common for dictators to create several *coercive institutions*, government agencies tasked with repressing dissent, with similar tasks and unclear authority hierarchies between them (Greitens, 2016). When dictators do this, they can use these parallel organisations to monitor each other. At a more general level, a dictator can make the military, a support party and secret police monitor each other (Geddes et al., 2018). This means that dictators can keep watch over the majority of their government apparatus

without the need for popular protest. Restrictions to freedom of assembly then become more advantageous, since the risks of popular protest remain while the benefits have decreased.

Succinctly, freedom of assembly offers a way to control the executive elite, but can be supplanted by using the elite to monitor itself once a dictator has sufficient power. My hypothesis on freedom of assembly is then:

H₂: *Higher levels of personalism are associated with lower levels of freedom of assembly in a country.*

Protection of life and physical integrity and freedom of movement: The competence of repressors

Protection of life and physical integrity is protection from forms of purposeful physical violence and violent coercion (e.g., torture), as well as protection against state-sanctioned killing (Coppedge et al., 2021a, p. 173; Davenport, 2007a, p. 2). *Freedom of movement* is the freedom of inhabitants of a country to go to and be at any location they wish to be for any amount of time (Coppedge et al., 2021a, pp. 182-184; Møller & Skaaning, 2013b, p. 1071; Universal Declaration of Human Rights, 1948, Art. 16).

Protection of life and physical integrity, and to a lesser extent freedom of movement, could deteriorate as personalism undermines the competence of *coercive institutions*, government agencies tasked with repressing dissent (Egorov & Sonin, 2020; Greitens, 2016). Less competent coercive institutions resort to means of repressing dissent that limit the civil liberties of large groups more readily (Greitens, 2016).

The reduction in the competence of coercive institutions follows from the clientelist divide-and-rule strategy that a personalising dictator uses, which can be illustrated with four tactics fitting such a strategy. First, rewards and job security based on loyalty might lead to neglect of competence as a selection criterion, greater tolerance of corruption, and reduced willingness to critique ineffectual policies among agents in coercive institutions (Svolik, 2012). Second, dictators might purposefully seek out less competent agents to prevent being outsmarted and removed (Egorov & Sonin, 2020; Libman, 2020). Third, dictators can use position rotation systems to control their elite, and officials that are constantly rotated can neither gain experience with local conditions nor nurture loyalty from their subordinates (Greitens, 2016). Fourth, when dictators create competing agencies fulfilling the same tasks to stimulate competition among them, the information gathered by these agencies remains fragmented and the agencies hamper each other's actions (Greitens, 2016). Overall, a dictator's strategy of divide and rule means the people selected into coercive institutions are

less capable that they would be under less personalised government, and have limited opportunities to accrue competence over time.

When the competence of coercive institutions decreases, they are more likely to resort to violent and indiscriminate measures that limit civil liberties on a large scale (Greitens, 2016). This happens because disorganisation within agencies, a lack of trust among agencies and little investment in informant networks among the general population limits insight into nascent threats (Egorov & Sonin, 2020; Greitens, 2016). And for the times when coercive institutions are able to detect a threat to the government, their persuasive power is limited because confused organisation means that they have little foothold among the general population (Greitens, 2016). With little insight into threats before they grow large and reduced means to control them in a peaceful manner, coercive institutions are more likely to have to resort to violence when a threat to the government is detected (Greitens, 2016). And as coercive institutions become less competent at assessing threats, their lack of insight also reduces their ability to deploy targeted measures against these threats (Greitens, 2016). Instead of violence against clear targets, coercive institutions then have to be more indiscriminate (Greitens, 2016). Inability to gather information on threats and identify targets of repressive action also means that coercive institutions are more likely to resort to large-scale monitoring that restricts freedom of movement. To illustrate, instead of singling out dissenters through intelligence-gathering, coercive institutions with less access to information might have to establish extensive access regulations and checkpoint systems to protect important locations or disrupt open resistance. Because of this, protection of life and physical integrity is probably violated more intensely and more indiscriminately under more personalised governments, and freedom of movement is reduced.

Because I expect personalism to gradually reduce both the protection of life and physical integrity and freedom of movement through the declining competence of coercive institutions, my hypotheses for these two civil liberties are:

- H3:** *Higher levels of personalism are associated to lower levels of protection of life and physical integrity in a country.*
- H4:** *Higher levels of personalism are associated to lower levels of freedom of movement in a country.*

Methods:

This section will start with a description of the dataset and peculiarities in the data collection. I will then discuss the operationalisations of concepts introduced in the theory section. This is supplemented with a description of my control variables and their operationalisation. The section ends with an analysis plan.

Dataset

In order to measure both personalisation and civil liberties, I combine two datasets. For personalisation, I use data on autocratic regimes from Geddes et al. (2018) extended by a latent variable model for personalism by Wright (2021). I retrieve civil liberties indicators from version 11.1 of the V-Dem dataset (Coppedge et al., 2021b), which uses latent variable models to provide continuous measures of various kinds of civil liberties. To obtain a regime type variable, I use another dataset by Geddes et al. (2014) which contains the same years and countries as the 2018 dataset. The Geddes et al. (2014, 2018) datasets run from 1946 to 2010 and cover autocratic regimes. V-Dem extends from 1789 to 2020 and also includes democratic regimes, which means that I only use a subset of the V-Dem dataset. All three datasets are time-series cross-sectional (or panel) data sets, containing yearly observations per country (country-years). Since years within a country will be related, the dataset has a multilevel structure with country-years at level 1 nested in countries over time in level 2. The full dataset includes 4591 country-years from 119 countries observed for 1 to 65 years, but the effective size when accounting for missingness is 3406 country-years from 109 countries observed for 1 to 55 years. I will report results for this subset.

Data collection procedure

Both the Geddes et al. (2014, 2018) datasets and the V-Dem dataset (Coppedge et al., 2021c) are coded by either the scholars constructing the dataset, expert surveys or research assistants. My subset of the V-Dem data also includes a number of variables extracted from other datasets (Coppedge et al., 2021c). This means that for many variables, the final values will depend on the availability of experts and their judgements on the state of a variable in a specific country-year. V-Dem uses multiple coders per country-year, and when feasible provides estimates of the uncertainty of reported values (Coppedge et al., 2021c). Geddes et al. (2014, 2018) do not provide extensive information about the number of coders or the reliability of their estimates.

Variable modelling

The values of personalism and the scale items for civil liberties have been created via Item-Response Theory (IRT) models (Pemstein et al., 2021; Wright, 2021). This means that the responses to basic indicators (as given by coders) were taken as imperfect measurement of a latent variable, and a probable value for this latent variable was inferred from aggregating the basic indicators (Pemstein et al., 2021). For the personalism score, this aggregation was done across a number of items (Wright, 2021), while the civil liberties measures aggregate across multiple coders' responses (Pemstein et al., 2021). For personalism, I use Wright's (2021) measure directly, while for the civil liberties I use a combination of direct scores, results from Bayesian factor analyses included in the V-Dem dataset (Pemstein et al., 2021, pp. 25-26) and simple sums and/or means (uniform weighing of items) calculated either by myself or by the V-Dem team.

Lags and coding dates

The Geddes et al. (2014, 2018) datasets use January 1st as the date of coding, while V-Dem measures averages over a year. If personalism increases in the middle of one year, these changes will only be reflected in the data in the next year. This means that V-Dem scores often reflect the effect of a change in the level of personalism in the previous year. An important advantage of this is that changes in personalism in a year then always precede changes in respect for civil liberties, satisfying one condition for causal inference. As such, the personalism scores are matched to their nominal years. Previous research on the relation between repression and personalism also used this approach (Frantz et al., 2019).

Operationalisations of core variables

My main variables of interest are personalisation and four indicators for various kinds of civil liberties. The civil liberties measures are measures of actual respect for or violation of civil liberties.

Personalism

The basic indicators with which personalism was estimated are eight binary variables recording whether access to high office depends on loyalty to the regime leader, a leader creates a new political support party after gaining power, a regime leader controls appointments to a party executive committee, the party executive committee is either absent or powerless (a ‘rubber stamp’), the leader personally controls the security apparatus, the regime leader promotes officers loyal to himself or his group and forces officers from other groups out of the military, the leader creates paramilitary forces and whether the leader imprisons or kills officers from other groups without a fair trial (Geddes et al., 2018; Wright, 2021).

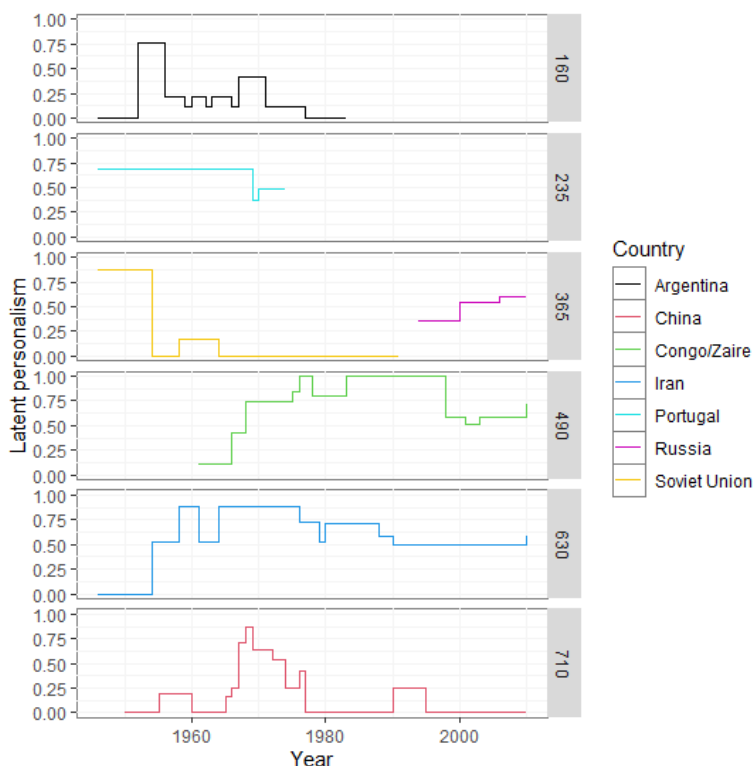
Wright (2021) aggregates these

indicators to a variable on the 0 to 1 range via an IRT model, with values closer to 1 indicating a greater degree of personalism. The indicators had Cronbach’s Alpha values between 0,66 and 0,83 in individual years. This variable was observed for all country-years in the full dataset ($N = 4591$, $n = 119$, $T = 1-65$).

A practical problem with a constructed variable to measure personalism is the lack of natural reference values for interpreting the variable. To aid interpretation of the personalism scores, I have added example time-series of personalism levels for six countries in Figure 1.

Figure 2

Time series of personalism in the People’s Republic of China, the Democratic Republic of Congo (and former Zaire), Russia (and the former Soviet Union), Portugal, Iran and Argentina



Note. The time series only depict the period in which a country was classified as autocratic by Geddes et al. (2018). The 1992-1993 break in the time series for the Soviet Union and Russia signifies the transition from one country to the other and a short period where Russia was deemed democratic in the Geddes et al. (2018) coding scheme.

To discuss a few examples, personalism rose to near unity for Mobutu Sese Seko, who is broadly recognised as a dictator with an exceptionally complete grasp on power in the Democratic Republic of the Congo/Zaire (e.g., Turner & Young, 1985). Meanwhile, communist autocracies such as the USSR and the People's Republic of China tend to have very low levels of personalism. Where this pattern is disrupted that is theoretically sensible, with episodes such as the dominant position of Joseph Stalin (McDermott, 2014) and the attempt to achieve greater power by Mao Zedong during the cultural revolution (Wright, 2021, pp. 5-6).

Freedom of expression

Freedom of expression is a variable combining IRT model values based on ordinal expert-coded assessments of the direct or indirect government censorship of print and/or broadcast media, harassment of journalists, media self-censorship, the freedom of ordinary people to discuss political matters at home and in the public sphere, and the freedom of academic and cultural expression (Coppedge et al., 2021a, p. 307). The indicators are aggregated by (Bayesian) factor analysis (Coppedge et al., 2021a, p. 307; Pemstein et al., 2020, pp. 25-26). Cronbach's Alpha varied between 0,80 and 0,93 for individual years. The variable runs from 0 to 1 with higher values indicating higher levels of freedom of expression. To avoid problems with floating point arithmetic when estimating models of freedom of expression, I multiplied the 0 to 1 scale a hundredfold (to a 0 to 100 scale).¹ The variable was observed for all country-years in the full dataset ($N = 4591$, $n = 119$, $T = 1-65$).

Freedom of assembly

Freedom of association and assembly is also measured by IRT model values based on ordinal expert coding ranging from 0 if state authorities do not allow peaceful assemblies and are willing to use lethal force to prevent them to 4 if state authorities almost always allow and actively protect peaceful assemblies except in rare cases of lawful, necessary and proportionate limitations (Coppedge et al., 2021a, p. 225). The resulting variable has a mean of zero in the complete V-Dem dataset (which includes democracies excluded from my dataset), with higher values indicating greater freedom of assembly. This variable had 4,40% of values missing ($N = 4389$, $n = 114$, $T = 1-65$).

¹ The personalism variable did not require this treatment because it is not a dependent variable. The models presented later were also estimated with personalism multiplied by a hundred, but this did not change conclusions.

Protection of life and physical integrity

Protection of life and physical integrity is measured as the mean of expert-coded ordinal indicators of freedom from torture and freedom from political killings rescaled to a 0 to 1 range (Coppedge et al., 2021a, p. 293) Higher values indicate greater protection of life and physical integrity. As with freedom of expression, I multiplied the scale by a hundred to avoid problems with floating point arithmetic. The items have correlations between 0,74 and 0,88 within individual years. The variable was observed for all country-years in the full dataset ($N = 4591$, $n = 119$, $T = 1-65$).

Freedom of movement

For freedom of movement, I averaged IRT model values from expert-coded ordinal indicators of freedom of domestic movement for men, freedom of domestic movement for women and freedom of foreign movement (Coppedge et al., 2021a, pp. 182-184). All three items have a mean of zero in the complete V-Dem dataset, with higher scores indicating greater freedom of movement. Cronbach's Alpha varies from 0,76 to 0,91 for individual years. The scale was computed as the mean of the three indicators. This variable was observed for all country-years in the full dataset ($N = 4591$, $n = 119$, $T = 1-65$).

Rigour and impartiality public administration

The rigour and impartiality of the public administration is measured by IRT model values based on expert coding ranging from 0 if the law is not respected by public officials and arbitrary or biased administration of the law is rampant to 4 if the law is generally fully respected by public officials and arbitrary or biased administration of the law is very limited (Coppedge et al., 2021a, pp. 175-176). The original (ordinal) variable was also coded 0 if no functioning public administration existed. Again, the resulting variable has a mean of zero for the complete V-Dem dataset with higher values indicating greater rigour and impartiality of the public administration. This variable was observed for all country-years in the full dataset ($N = 4591$, $n = 119$, $T = 1-65$).

Control variables and operationalisations

To avoid bias in effect estimates from my observational data, I include a number of control variables that might change over time and possibly between countries and over time. I believe that these variables can affect both respect for civil liberties and personalism, but I will only sketch a possible mechanism for each variable.

Regime type

Regime type is a nominal variable with four categories: single-party regimes, military regimes, personalist regimes and monarchies. These categories identify whether control over policy, leadership selection and the security apparatus in an autocratic regime are primarily in the hands of a ruling party (single-party), the military (military), a narrower group centred around an individual dictator (personalist) or a royal family (monarchy, Geddes et al., 2014, p. 318). Wright (2021, pp. 7-9) shows that these regime type variables are distinct from the personalism variable, which will be reflected in my bivariate results presented later. The variable is constructed from a set of four mutually exclusive (linearly dependent) binary variables, of which three are entered into regressions (Geddes et al., 2014).

International or internal conflict

The most important control variables are separate binary variables indicating whether the country was in international conflict or internal conflict (Coppedge et al., 2021a, pp. 362-363), which previous work has shown to be an important predictor of respect for civil liberties (Hill & Jones, 2014). Conflict might also affect personalism, for example because a government under external pressure is more likely to concentrate formal responsibilities in a single individual (e.g., Geddes et al., 2018, pp. 160-162), or conflict forces military spending that creates a powerbase independent from the dictator (e.g., Geddes et al. 2018, pp. 162-163; Svobik, 2012, pp. 127-133). These variables were always missing simultaneously, which was the case for 16,99% of all country-years ($N = 3811$, $n = 115$, $T = 1-55$).

Political violence

Since political violence is a well-established predictor of government repression (e.g., Davenport et al., 2021) and political violence might accompany or encourage personalisation, I include a measure of political violence by non-state actors (Coppedge et al., 2021a, pp. 224-225). This variable had 4,47% of its values missing ($N = 4386$, $n = 114$, $T = 1-65$).

Population size

I control for logged (\log_{10}) population size (Coppedge et al., 2021a, p. 360), as population size might affect the functioning of government both with regard to its tendency to respect civil liberties and the degree of personalism. Logged population size had 18,62% missing values ($N = 3736$, $n = 113$, $T = 1-55$).

GDP per capita and GDP growth

I include both GDP growth and the logged (ln) GDP per capita (Coppedge et al., 2021a, p. 358-359), as both the level of economic prosperity and changes in it may affect the political climate such that they might make personalism more (or less) likely, but also make restrictions on civil liberties more (or less) likely. Logged GDP per capita had 5,97% of values missing ($N = 4317$, $n = 114$, $T = 1-61$). For GDP growth, the percentage of missing values was 6,32% ($N = 4301$, $n = 114$, $T = 1-61$).

Analysis plan

Since I will be estimating models for four kinds of civil liberties, and thus running a large number of independent statistical tests, I will apply a Bonferroni correction to my significance level of $\alpha = 0,05$ and use a significance level of $\alpha = 0,0125$ at the level of a single test. I will carry out my analyses in R (R Core Team, 2021) using the “lme4” package to estimate multi-level models (Bates et al., 2015) extended by the “clubSandwich” package for clustered standard errors (Pustejovsky & Tipton, 2016). I estimate all models with REML, but re-estimate using normal ML when comparing models.

The models I will be estimating are linear random intercept models of the form:

$$CL_{i,t} = \alpha_i + \lambda_t + \beta_0 + \beta_P P_{i,t} + \beta_{RI} RI_{i,t} + \sum_{k=1}^K \beta_k X_{k,i,t-1} + \sum_{j=1}^3 \beta_{K+j} RT_{j,i,t} + \varepsilon_{i,t}$$

where I assume that countries have persistent differences in their mean level of respect for civil liberty $CL_{i,t}$ over time, modelled by deviations from a between-country intercept β_0 using the term α_i . These country-specific deviations have a normal distribution with a mean of 0 and a standard deviation σ_α . Note that these country deviations can be regarded as an error term, as they use those parts of deviations of observed values from predicted values that are stable over time.

In addition to the random country effects, I model global changes in the level of respect for a civil liberty over time by year-fixed effects λ_t . These are dummies for all years except 1946 (which is represented by the intercept). I assume that the effect β_P of personalism ($P_{i,t}$) is the same in every country i . Rigour and impartiality of the public administration ($RI_{i,t}$) and the control variables $X_{k,i,t-1}$ are also assumed to have the same effect (β_{RI} , β_k) in every country. To account for the fact that personalism is coded to the value on the first of January of year t , I lag the control variables by one year (to $t - 1$). The regime type variables with personalist

regimes as the reference category ($RT_{j,i,t}$) are an exception, as these also have the first of January of year t as coding date.

While the random country intercepts can be regarded as a level 2 error term, the model also has a level 1 error term ($\varepsilon_{i,t}$) for the part of deviations from predicted values that changes over time. I assume these errors to follow a normal distribution with a mean of 0 and a different standard deviation $\sigma_{\varepsilon,i}$ in each country using Pustejovsky and Tipton's (2017) extension of MacKinnon and White's (1985) heteroscedasticity-robust standard errors. This means that the variances used to calculate the standard errors differ from the within-country variance reported for my models, since this within-country variance is estimated assuming homoscedasticity. The random country intercepts and standard error clustering correct for the dependence of prediction errors in the same country that result from persistent differences between countries, but not for dependence in prediction errors caused by the tendency for values in the same country to be more similar as they are closer in time.

For the civil liberties where I do not assume a mediation through the rigour and impartiality of the public administration, I will first estimate the random intercept model without year-fixed effects or predictors. In a second model I add the year-fixed effects. A third model adds the personalism variable. Then, I estimate a fourth model where I add the control variables (international conflict, internal conflict, political violence, population size, GDP per capita, GDP growth, and three regime type dummies). For the civil liberties where I do not assume a mediation, this fourth model is used to test my hypothesis of personalism being related to the civil liberty in question. The other models clarify this model by showing the stability of the relation to control variables and showing the distribution of variance between countries and within countries. In the case of freedom of expression, I also estimate a model with personalism and its square as predictors in addition to year-fixed effects to test my hypothesis of a curvilinear association. If this model suggests a curvilinear association between personalism and freedom of expression, I will incorporate that into the final model.

For the civil liberties where I can test mediation, I further estimate a fifth model where I not only include the control variables, but also the rigour and impartiality of the public administration. In these cases, this model tests my mediation hypothesis. In addition to the models with civil liberties as the dependent variable, I will estimate random intercept models with the rigour and impartiality of the public administration as dependent variable. Here, I first estimate a model with random country effects, add year-fixed effects in a second model, estimate a third model with personalism as predictor and add the control variables also used in the other regressions as a fourth step (international conflict, internal conflict, political

violence, population size, GDP per capita, GDP growth, and three regime type dummies). Similar to the models for civil liberties without mediation by the rigour and impartiality of the public administration, I use the last model to test my hypothesis that personalism influences the rigour and impartiality of the public administration while the other models contextualise and clarify this model.

Results:

In this section, I will present the results of my analyses on the dataset introduced in the last section. I will first give univariate and bivariate descriptives of the dataset to aid interpretation of further results, and then move on to a discussion of the estimated models. R script for all analyses can be found in the Appendix.

Univariate statistics

To get an impression of univariate distributions of my variables, I will first discuss the distribution of over-time country means for the various variables. These give a first indication of the chances that the assumption of normally distributed country effects will be met and provide a succinct summary of the data, given that within-country observations have over-time dependence and are thus probably quite similar to each other and their mean. When describing the distribution of over-time country means I provide intra-country correlations, but to give more insight into the over-time developments in variables I will separately discuss some important over-time changes in the univariate distributions.

To start on the first point, the over-time country means are summarised by their mean, standard deviation, skew and kurtosis in Table 1. To get a first impression of missing data effects, these descriptive statistics are reported for both the full dataset and complete cases (where any country-year with a missing value on any of the variables is removed). On average, personalism is at a medium level of 0,40 among complete cases. There is a fair amount of variation around this mean, with a standard deviation of 0,21. The distribution of personalism is nearly symmetric ($Skew = 0,02$), and has fewer extreme values than a normally distributed variable ($Kurtosis = -0,67$). There is correlation in personalism within countries, but this intra-country correlation is actually not that large ($ICC = 0,41$). This suggests that the level of personalism in a given year tells something about the level of personalism in other years in the same country, but not so much as to have the level of personalism be very stable over long periods.

The mean of freedom of movement is not especially low or high ($M > -0,01$), that of freedom of assembly ($M = -0,90$) is somewhat low. Since these variables are constructed in such a way that their mean in the entire V-Dem dataset containing both autocracies and non-autocracies is zero, the mean level of freedom of assembly in autocracies is lower than the overall mean in both autocracies and non-autocracies. Given a theoretical 0-100 range, protection of life and physical integrity ($M = 35,16$) and freedom of expression ($M = 28,93$) are low. Between-country standard deviations are quite high for all civil liberties, as the range

of the variables can be covered by less than six standard deviations while a normally distributed variable would have 99,7% (instead of 100%) of its observations within three standard deviations from the mean. This ratio would suggest lower kurtosis than a normal distribution if the variables were symmetrically distributed. However, freedom of movement and freedom of expression both have above-normal kurtosis (1,25 and 0,53), which is possible because both variables are skewed (-0,86 and 0,96 respectively). Protection of life and physical integrity and freedom of assembly have below-normal kurtosis (-0,30 and -0,20 respectively), but while freedom of assembly is closer to symmetry ($Skew = 0,27$) protection of life and physical assembly has clear right-skew ($Skew = 0,72$). The intra-country correlation of the civil liberties varies somewhat, from freedom of movement having rather strong over-time dependence ($ICC = 0,76$) to freedom of expression varying considerably more within countries over time ($ICC = 0,46$).

For the control variable, it is noteworthy that the mean population level in the countries in the complete dataset over time is approximately 23.000.000 people ($10^{\overline{\log_{10}(Population)}} \approx 23.294.375$) and the over-time mean GDP per capita across all countries is approximately \$4.500 ($e^{\overline{\ln(GDP\ per\ capita)}} \approx \$4.423,96$) in 2011 US dollars (Bolt et al., 2018), while the average over-time mean GDP growth per year is 2% ($M = 0,02$). The average level of rigour and impartiality of the public administration in the autocratic countries in the complete dataset is somewhat below the average level in all (autocratic and non-autocratic) countries in the V-Dem dataset ($M = -0,60$). Population levels are very strongly related within one and the same country ($ICC = 0,90$), while the intra-country correlation of GDP growth is very low ($ICC = 0,02$). GDP growth also has a rather large number of average GDP growth levels far from the country mean ($Kurtosis = 1,82$).

Figure 3 shows the by-year distribution of personalism using a heatmap that bins observations by value-year combinations (like a bivariate histogram), and this suggests that very large outliers (of up to 272%) that distort the mean of the countries with those outliers might explain the high kurtosis. Lastly, the distribution of over-time country

Figure 3

Heatmap of GDP growth over time for complete cases

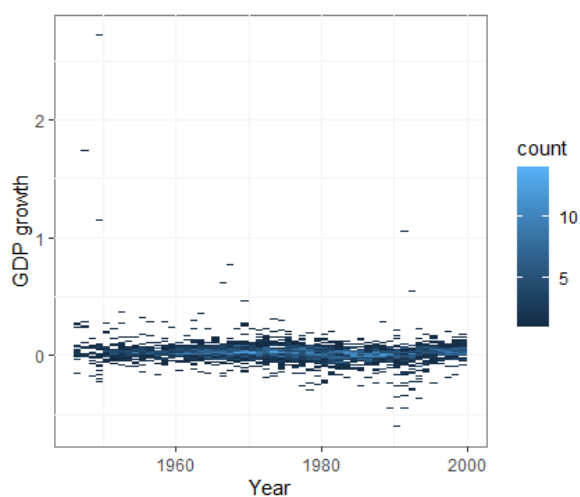


Table 1

Comparison of descriptive statistics between complete cases and the full dataset (M = mean, SD = standard deviation, S = skew, K = kurtosis, ICC = intra-country correlation)

Variable	Data	M	SD	S	K	ICC
Personalism	Full	0,41	0,21	-0,01	-0,55	0,52
	Complete	0,40	0,21	0,02	-0,67	0,55
Freedom of movement	Full	-0,01	1,08	-0,73	1,16	0,82
	Complete	>-0,01	1,13	-0,86	1,25	0,86
Freedom of expression	Full	29,28	18,26	0,87	0,48	0,60
	Complete	28,93	18,85	0,96	0,53	0,65
Freedom of assembly	Full	-0,91	1,01	0,19	-0,24	0,71
	Complete	-0,90	1,00	0,27	-0,20	0,71
Protection of life and physical integrity	Full	35,11	21,03	0,75	-0,15	0,71
	Complete	35,16	21,77	0,72	-0,30	0,75
Political violence	Full	-0,03	1,28	-0,05	-0,71	0,78
	Complete	-0,07	1,32	0,02	-0,70	0,81
log ₁₀ (Population)	Full	6,88	0,55	0,48	0,71	0,94
	Complete	6,89	0,55	0,44	0,68	0,94
ln(GDP per capita)	Full	8,06	0,82	0,40	-0,24	0,79
	Complete	7,99	0,81	0,29	-0,31	0,83
GDP growth	Full	0,02	0,02	0,49	0,20 ^a	0,04
	Complete	0,02	0,03	-0,24	1,82	0,04
Rigour and impartiality public administration	Full	-0,58	0,96	0,54	0,33	0,71
	Complete	-0,60	0,97	0,43	0,08	0,73

Note: Kurtosis is excess kurtosis, such that the normal distribution has a kurtosis of 0.

^a The comparatively low kurtosis of GDP growth for the full dataset results from a number of countries with outliers being dropped from the calculation due to ill-defined means.

means of political violence is notable because it has far lower kurtosis than a normally distributed variable (*Kurtosis* = -0,70).

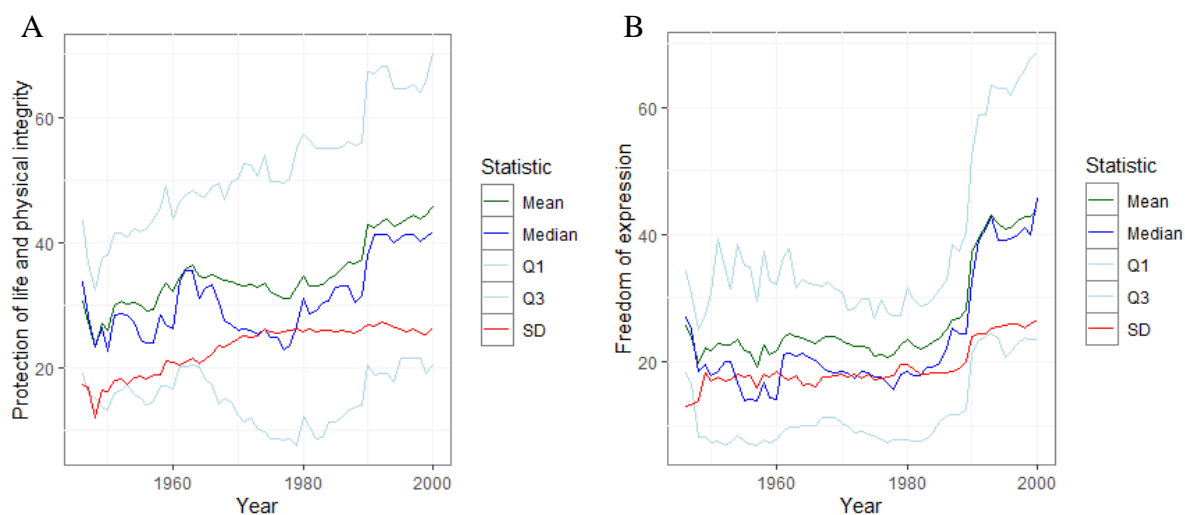
To expand on the intra-country correlations reported in Table 1, I will also describe two important over-time changes in the variables. The most important change in this regard is shown in Figure 4, which depicts how the yearly summary statistics of protection of life and physical integrity and freedom of expression evolve over time. Around the end of the Cold War in 1990, there is a clear upward shift in freedom of expression from approximately 0,2 to approximately 0,4. Protection of life and physical integrity also increases, but comparatively less. Generally, the civil liberties increase to varying extents around the end of the Cold War.

Interestingly, there is no clear shift in the proportion of single-party regimes around the end of the Cold War (See the Appendix, lines 2560 to 2606). A second point of interest is that the distribution of personalism moves steadily upward throughout the time period, starting at a mean of approximately 0,3 and ending up close to 0,5 by 2000.

As Table 1 suggests, the differences between the full dataset and complete cases are minor, with an important exception. Missingness of internal conflict and international conflict means that the complete cases only cover the period from 1946 to 2000 instead of running until 2010. This means that the data available cannot be used for conclusions about the early twenty-first century and developments therein.

Figure 4

Summary statistics of protection of life and physical integrity and freedom of expression over time for complete cases



Note: Panel A shows protection of life and physical integrity, Panel B shows freedom of expression. The mean is shown in dark green, the median in dark blue, the first and third quartile in light blue and the standard deviation in red.

Bivariate statistics

Although modelling will focus on variation within countries, accessing this variation without having intra-country correlation distort results is rather difficult without using a formal modelling approach. This section will only present the bivariate correlation between over-time country means (between-country correlations) in Table 2. I will focus on theoretically relevant correlations that are significant at the five percent level, which is any correlation of at least 0,19 for 109 independent observations.

Countries that have a higher average level of any one civil liberty also tend to have high values for other civil liberties ($0,44 \leq r \leq 0,78$). This means that the effect of personalism on any one civil liberty might be distorted when not controlling for the effect of other civil liberties. Beyond the correlations among the civil liberties, higher mean levels of personalism tend to be accompanied by lower average protection of life and physical integrity ($r = -0,24$), while the associations between personalism and the other civil liberties are statistically indistinguishable from zero. Greater rigour and impartiality of the public administration over time has a strong association with greater protection of life and physical integrity over time ($r = 0,61$). Countries with higher over-time means for the rigour and impartiality of the public administration also have more freedom of movement ($r = 0,36$), greater freedom of expression ($r = 0,44$), and greater freedom of assembly ($r = 0,38$). Although the association is in the expected direction, mean rigour and impartiality of the public administration is not significantly associated with mean levels of personalism ($r = -0,14$).

As expected, countries with more years under personalist regimes have higher average personalism ($r = 0,44$), although personalistic regimes and regimes with high levels of personalism are still statistically distinguishable. Countries with more years under military or single-party regimes are less personalistic ($r = -0,27$ and $r = -0,23$ respectively). The average levels of personalism in countries under a monarchy for differing lengths of time are statistically indistinguishable ($r = 0,06$). Beyond the regime type variables, other control variables do occasionally have sizeable correlations with civil liberties, but seldomly have a strong correlation to personalism. For example, political violence is strongly related to protection of life and physical integrity ($r = -0,52$), but too weakly related to personalism to say anything sensible about the relation ($r = -0,01$). This suggests that these control variables will probably not be particularly important in controlling the relation between personalism and civil liberties. At the same time, the between-country correlations in Table 2 might not accurately represent the within-country associations.

Table 2
Bivariate associations between over-time country means for complete cases

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. Personalism	-															
2. Freedom of movement	-0,13	-														
3. Freedom of expression	-0,18	0,72	-													
4. Freedom of assembly	-0,17	0,58	0,78	-												
5. Protection of life and physical integrity	-0,24	0,44	0,57	0,51	-											
6. Political violence	-0,01	-0,11	-0,15	-0,14	-0,52	-										
7. log ₁₀ (Population)	-0,15	-0,14	-0,13	-0,20	-0,22	0,21	-									
8. ln (GDP per capita)	-0,23	0,06	-0,02	-0,10	0,32	-0,27	-0,10	-								
9. GDP growth	-0,19	-0,05	-0,13	-0,12	0,14	-0,14	0,01	0,33	-							
10. Rigour and impartiality public administration	-0,14	0,36	0,44	0,38	0,61	-0,47	-0,20	0,18	0,25	-						
11. International conflict	-0,03	-0,04	-0,02	-0,02	-0,04	0,17	0,07	0,05	0,09	0,13	-					
12. Internal conflict	-0,01	-0,12	-0,13	-0,22	-0,30	0,42	0,24	-0,14	-0,04	-0,34	-0,02	-				
13. Personalist regime	0,44	0,26	0,16	0,08	-0,16	0,14	-0,06	-0,18	-0,35	-0,11	0,08	0,09	-			
14. Military regime	-0,27	0,02	-0,12	-0,15	-0,24	0,27	0,11	0,11	0,15	-0,26	-0,13	0,11	-0,22	-		
15. Single-party regime	-0,23	-0,20	<0,01	0,06	0,18	-0,17	0,09	-0,08	0,04	0,21	-0,02	-0,13	-0,60	-0,40	-	
16. Monarchy	0,06	-0,07	-0,10	-0,05	0,20	-0,23	-0,20	0,25	0,26	0,11	0,08	-0,04	-0,20	-0,17	-0,32	-

Note: n is 109, correlations larger than 0,19 are significant at $\alpha = 0,05$

Regression modelling

Freedom of expression

The first civil liberty for which I will discuss the regression models is freedom of expression. I will report the results for the first two models with random country intercepts and random country intercepts plus year-fixed effects in running text, while the models adding theoretically important predictors are also reported in Table 3.

Model 1 containing only the random country intercepts can be used to see which variation in freedom of expression is attributable to persistent between-country differences, and which variation lies in over-time changes within countries. The estimates suggests that most variation is between-country variance at 330,36, while pooled within-country variance is lower at 177,95. The variance decomposition implies an intra-country correlation of 0,65, while the between-country intercept is 28,81 ($SE = 1,77$, 98,75% CI [24,33; 33,30]). Since Model 1 does not contain any regressors, these statistics should and do approximately match the statistics reported in Table 1 ($M = 28,93$, $ICC = 0,65$).

Adding the year-fixed effects in Model 2 significantly improves the model fit ($\chi^2(54) = 789,67$, $p < 0,001$). This is also reflected in the decrease in both between-country variance and within-country variance from Model 1 to Model 2. There is a proportional decrease by 12,24 percent in variance, calculated by taking subtracting the total variance in Model 2 compared to that in Model 1 from one ($R_1^2 = 1 - (\hat{\sigma}_{\varepsilon,2} + \hat{\sigma}_{\alpha,2})/(\hat{\sigma}_{\varepsilon,1} + \hat{\sigma}_{\alpha,1})$; Snijders & Bosker, 2012, p. 112). The intra-country correlation is a little higher than in Model 1 at 0,68. The mean level of freedom of expression rises significantly after 1990, with the 1946 mean level being 27,61 ($SE = 3,64$, 98,75% CI [17,59; 37,62]) while the 2000 mean level is $27,61 + 15,26 \approx 42,86$ ($b = 15,26$, $SE = 4,51$, 98,75% CI [2,82; 27,69]).

Model 3 is the first model reported in Table 3, and adds the personalism variable. Countries with the highest possible level of personalism are predicted to have a level of freedom of expression that is 17,03 points lower ($SE = 3,46$, 98,75% CI [-26,28; -9,94]) than countries with the lowest possible level of personalism, corrected for the mean level of freedom of expression in that country over time and the mean level of freedom of expression across countries in a year. I deem this a small-to-medium sized effect, since a country that has the lowest possible level of personalism instead of the highest possible level would in expectation have only about a within-country standard deviation's lower level of freedom of assembly. Model 3 significantly improves model fit compared to Model 2 ($\chi^2(1) = 239,68$, $p < 0,001$). The year-fixed effects still improve the model fit, given that personalism is in the

model ($\chi^2(54) = 898,47, p < 0,001$). The decrease in total compared to Model 1 is 17,85 percent, and since the intra-country correlation is still approximately 0,68 between-country and within-country variance decreased roughly equally.

Model 4 adds the square of personalism to the model with just a linear personalism term. This squared term only minimally decreases total variance compared to Model 3 ($R_I^2 = 0,19$), while neither the linear personalism term ($b = -7,29, SE = 8,22, 98,75\% \text{ CI } [-28,62; 14,04]$) nor the quadratic term ($b = -11,20, SE = 9,62, 98,75\% \text{ CI } [-36,36; 13,96]$) is significant. Since the differences with Model 3 are minimal and the quadratic term is insignificant, my hypothesis of a curvilinear association between personalism and freedom of expression seems implausible. The residual analysis of Model 5 will provide yet more evidence against the hypothesis of curvilinearity.

Model 5 adds political violence, logged population, logged GDP per capita, GDP growth, internal and international conflict and regime type as control variables. The effect of personalism changes little with the addition of the control variables ($b = -18,11, SE = 3,14, 98,75\% \text{ CI } [-25,38; -8,71]$), while none of the control variables is significant at $\alpha = 0,0125$. The control variables do seem to improve model fit together compared to Model 3 ($\chi^2(9) = 177,73, p < 0,001$), but between-country variance increases raising the intra-country correlation to 0,72. The total decrease in variance compared to Model 1 is 10,71 percent, less than in Model 3 to such an extent that the model might be misspecified (Snijders & Bosker, 2012, p. 156). The year-fixed effects still contribute to model fit ($\chi^2(54) = 718,45, p < 0,001$).

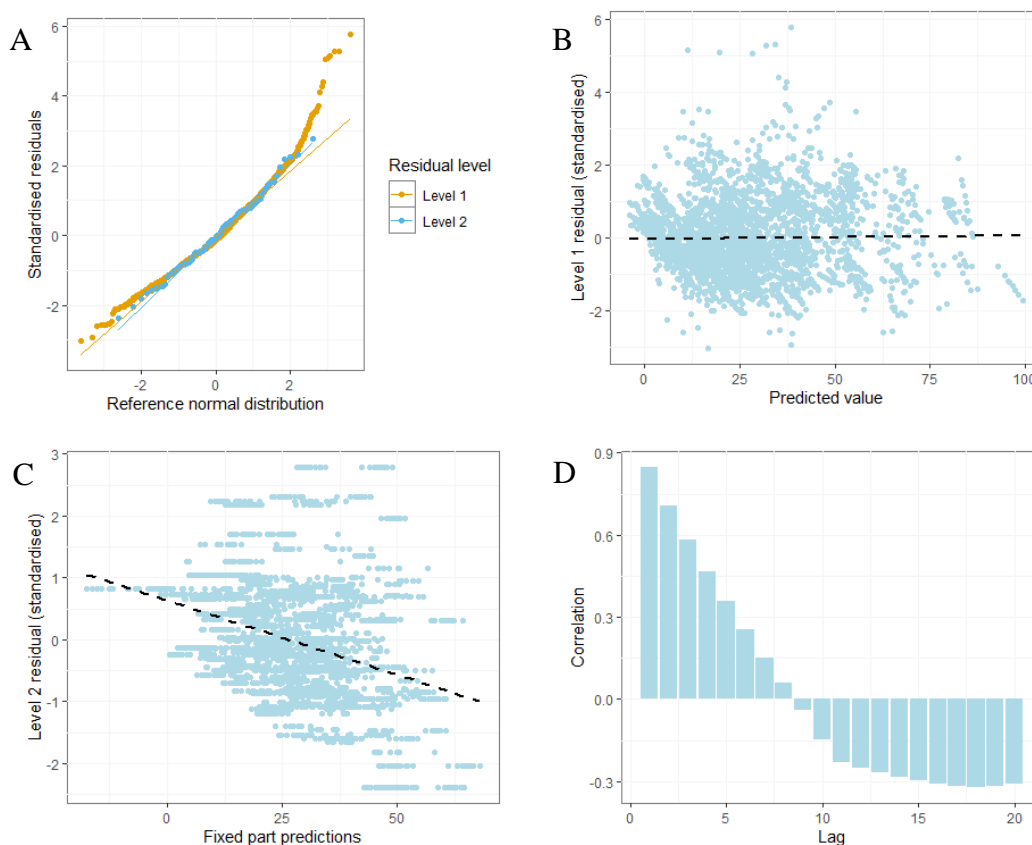
I estimated a model adding the rigour and impartiality of the public administration (lines 3585 to 3646 in the Appendix) to test my assumption that this would not be a meaningful mediating variable for freedom of expression. My assumption seems questionable, as the effect of personalism in this model is decreased ($b = -11,38, SE = 3,02, 98,75\% \text{ CI } [-19,23; -3,52]$) while the rigour and impartiality of the public administration has a strong association with freedom of expression ($b = 10,00, SE = 1,33, 98,75\% \text{ CI } [6,52; 13,47]$).

To assess the validity of Model 4, I analysed the residuals of the model. The residuals can be separated into two ‘levels’. The level 2 residuals are differences between general intercept and the country-specific intercepts, showing the distribution of the random intercept term. The level 1 residuals stem from differences between predicted values (including random effects) and observed values. The level 1 residuals are intuitively more similar to the residuals of a ‘normal’ single-level model. The diagnostic plot for the level 1 residuals shown in Panel B of Figure 5 suggest that a straight-line approximation of the relation between

freedom of expression and the predictors (including personalism) does not lead to systematic errors in my predictions. This conflicts with the hypothesis of a curvilinear association. However, the residuals are right-skewed as shown in Panel A, which means that my assumption of normally distributed errors used to calculate confidence intervals is somewhat questionable. Panel C shows that standardised level 2 residuals are related to predicted values, which suggests that persistent differences in freedom of expression are related to one or more predictors. Panel D shows strong residual autocorrelation for the level 1 residual, which means that prediction errors for the same country in one year and subsequent years are very similar, violating the assumption of independent errors.

Figure 5

QQ-plot, predicted value-standardised residual plots and autocorrelation plot for freedom of expression.



Note: The standardised residuals are standardised with country-level standard

Table 3

Linear random intercept model with freedom of expression as dependent variable and year-fixed effects (N = 3406, n = 109, T = 1-55).

Variable	Model 3		Model 4		Model 5	
	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>
Intercept	33,36 (3,30)	<0,01	26,39 (2,91)	<0,01	127,16 (49,86)	0,01
Personalism	-17,03 (3,46)	<0,01	-7,29 (8,22)	0,38	-18,11 (3,14)	<0,01
Personalism ²			-11,20 (9,62)	0,25		
Political Violence					0,70 (1,29)	0,59
log ₁₀ (Population)					-13,46 (7,15)	0,07
ln(GDP per cap)					-0,42 (2,39)	0,86
GDP Growth					-2,39 (2,10)	0,28
Internal conflict (1 = conflict)					-2,13 (1,60)	0,19
International conflict (1 = conflict)					-1,72 (1,44)	0,25
Regime type						
Personalist (ref)						
Military					-2,64 (3,27)	0,43
Single-party					-5,86 (4,74)	0,22
Monarchy					9,98 (6,43)	0,16
Variance						
Between- country	284,46		280,89		327,97	
Within- country	133,11		132,69		125,90	
<i>R</i> _l ² ^b	0,18		0,19		0,11	
<i>ICC</i>	0,68		0,68		0,72	
<i>LR</i> χ^2 Year-fixed ^c	898,47	<0,01	898,49	<0,01	697,50	<0,01
<i>LR</i> χ^2 ^d	239,68	<0,01	12,92	<0,01	177,73	<0,01

Note: ^a The standard errors are country-clustered. ^b Proportional decrease in level 1 variance compared to Model 1. ^c Compares the model with year effects to an equivalent model without them. ^d Compares Model 3 to Model 2, Model 4 to Model 3 and Model 5 to Model 3.

Freedom of assembly

As with freedom of expression, the first models are only reported in running text. Model 1 includes only the random country intercepts. Between-country variance is 0,96 and within-country variance is 0,39. Most of the variance in freedom of assembly is thus located between countries, making freedom of assembly fairly stable over time within countries (*ICC* = 0,71). The general mean of freedom of assembly is -0,90 (*SE* = 0,09, 98,75% CI [-1,14; -0,66]), as

is to be expected from Table 1. Model 2 adds year-fixed effects to the random country intercepts, which significantly improves model performance ($\chi^2(54) = 519,49, p < 0,001$) although the individual comparisons between 1946 ($b = -0,76, SE = 0,19, 98,75\% \text{ CI } [-1,29; -0,24]$) and later years rarely point to individually significant differences. Unexplained variance is lowered by taking changes over time into account ($R_I^2 = 0,07$), but the ratio of between-country to within-country variation is more or less the same ($ICC = 0,73$).

Model 3 adds personalism as a continuous predictor, and this significantly improves the model fit ($\chi^2(1) = 156,10, p < 0,001$) with a total variance decrease of 11,14 percent compared to Model 1. Countries at the highest level of personalism tend to have a value for freedom of assembly that is 0,67 points lower ($SE = 0,17, 98,75\% \text{ CI } [-1,12; -0,22]$) than that of countries at the lowest level of personalism, given the long-term mean of the country and the mean level of freedom of assembly in a year. Again, this seems like a small-to-medium-sized effect. Model 4 controls the effect of personalism for a selection of control variables in addition to a country's persistent deviation from the cross-country mean level of freedom of assembly and year-fixed effects. These control variables seem to significantly improve model fit together ($\chi^2(9) = 248,51, p < 0,001$). However, none of the individual coefficients differs significantly from zero and the variance is substantially greater than that of Models 2 or 3 ($R_I^2 = 0,02$). The relation between freedom of assembly and personalism is somewhat stronger in Model 4 than in Model 3 ($b = -0,71, SE = 0,16, 98,75\% \text{ CI } [-1,12; -0,30]$), but the difference is not statistically or substantively meaningful. Models 3 and 4 fit the hypothesised decrease in freedom of assembly with greater personalism.

As with freedom of expression, I also estimated a model including the rigour and impartiality of the public administration for freedom of assembly (See lines 3951 to 4010 in the Appendix). Again, the rigour and impartiality of the public administration has a statistically significant association with freedom of assembly ($b = 0,42, SE = 0,06, 98,75\% \text{ CI } [0,27; 0,57]$) while adding that variable makes the association of personalism to freedom of assembly substantially weaker ($b = -0,43, SE = 0,15, 98,75\% \text{ CI } [-0,81; -0,05]$).

The most important parts of the residual analysis for Model 4 of freedom of assembly are shown in Figure 6. Panel A shows that the level 2 residuals are fairly normally distributed. The level 1 residuals have a rather stretched right tail, although the distribution is fairly close to a normal distribution otherwise. Panel B suggests that residual values still have fairly strong persistence over time. Further predicted value-residual plots of the level 1 and level 2 residuals showed the level 1 residuals to conform to the linear model fairly well while level 2 residuals have a clear downward slope.

Table 4

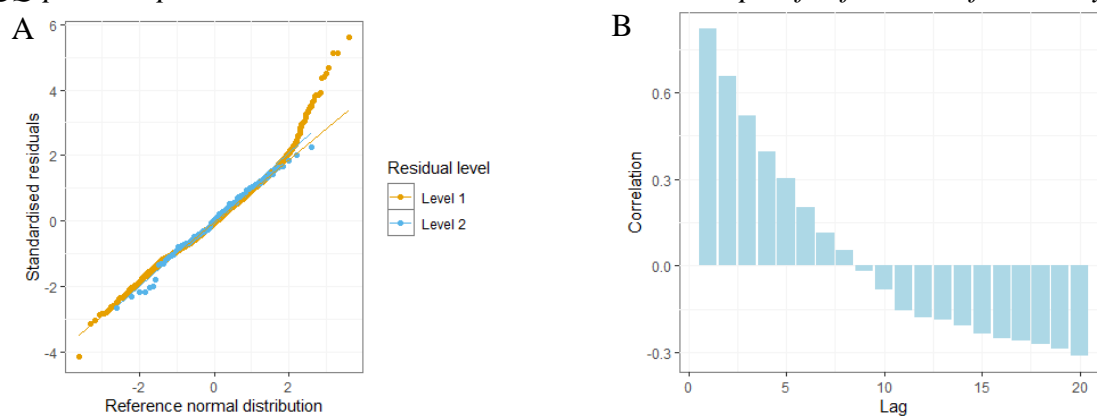
Linear random intercept model with freedom of assembly as dependent variable and year-fixed effects (N = 3406, n = 109, T = 1-55).

Variable	Model 3		Model 4	
	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>
Intercept	-0,54 (0,20)	0,01	4,69 (2,92)	0,11
Personalism	-0,67 (0,17)	<0,01	-0,71 (0,16)	<0,01
Political Violence			-0,12 (0,07)	0,08
log ₁₀ (Population)			-0,80 (0,41)	0,06
ln(GDP per cap)			0,01 (0,12)	0,94
GDP Growth			-0,08 (0,12)	0,54
Internal conflict (1 = conflict)			0,06 (0,09)	0,52
International conflict (1= conflict)			0,04 (0,10)	0,73
Regime type				
Personalist (ref)				
Military			-0,21 (0,15)	0,17
Single-party			-0,20 (0,19)	0,31
Monarchy			0,78 (0,31)	0,04
Variance				
Between-country	0,87		1,01	
Within-country	0,32		0,30	
R_I^2 ^c	0,11		0,03	
ICC	0,73		0,77	
$LR \chi^2$ Year-fixed ^b	562,54	<0,01	570,00	<0,01
$LR \chi^2$ ^d	156,10	<0,01	248,51	<0,01

Note: ^a Country-clustered standard errors. ^b Compares model with year effects to an equivalent model without them. ^c Proportional decrease in level 1 variance compared to Model 1. ^d Compares each model to the previous model.

Figure 6

QQ-plot and predicted value-standardised residual scatterplot for freedom of assembly.



Note: The standardised residuals are standardised with country-level standard deviations.

Protection of life and physical integrity

For protection of life and physical integrity, Model 1 including only random country intercepts estimates between-country variance to be 460,35 while within-country variance is 153,77. This suggests that variation in protection of life and physical integrity mostly occurs between countries ($ICC = 0,75$). The general mean for protection of life and physical integrity is 35,25 ($SE = 2,07$, 98,75% CI [29,98; 40,52]), close to that in Table 1. While the year-fixed effects added in Model 2 significantly improve model fit ($\chi^2(54) = 199,63$, $p < 0,001$), individual differences compared to 1947 ($b = 39,16$, $SE = 4,58$, 98,75% CI [26,72; 51,60]) are not significant. Both within-country variance (147,23) and between-country variance (452,32) are a little lower in Model 2 than in Model 1 ($R^2 = 0,02$). These variances give an intra-class correlation of 0,75.

Model 3 shows that countries with the highest possible level of personalism have a level of protection of life and physical integrity that is 17,48 points lower on average ($SE = 3,94$, 98,75% CI [-27,70; -7,26]) than the level in countries with the lowest possible level of personalism, given a country's persistent influences on protection of life and physical integrity and the mean level of protection of life and physical integrity in a year across countries. This effect is somewhat small, taking into account that the estimated coefficient represents the largest difference personalism can be associated with. Model 4 introduces control variables into the specification, and these control variables together significantly contribute to improving model fit ($\chi^2(9) = 481,15$, $p < 0,001$). The relation between personalism and protection of life and physical integrity in the Model 4 remains fairly similar to that in Model 3 ($b = -16,96$, $SE = 3,47$, 98,75% CI [-25,97; -7,95]), now controlling for the control variables as well as across-country year means and country effects. Overall, this is in line with my hypothesis that greater personalism is associated to less protection of life and physical integrity.

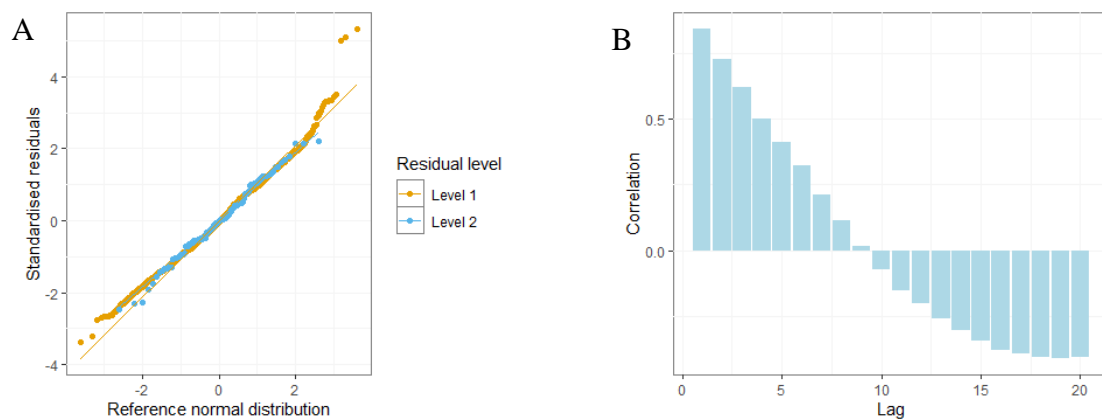
Model 5 adds the rigour and impartiality of the public administration. Greater levels of rigour and impartiality of the public administration tend to be accompanied by greater protection of life and physical integrity ($b = 11,23$, $SE = 1,53$, 98,75% CI [7,23; 15,24]) controlling for the level of personalism, the values for the control variables in the previous year, time-persistent differences between countries and the mean level of protection of life and physical integrity in a year. This is a rather large effect, as a shift from the minimum to the maximum of the range for the rigour and impartiality of the public administration (as with personalism) would amount to a shift of 74,95 points on the 100-point theoretical range in protection of life and physical integrity. The relation between personalism and protection of

life and physical integrity is almost halved ($b = -9,33$, $SE = 2,78$, 98,75% CI [-16,54; -2,11]), suggesting that this relation runs in part through the rigour and impartiality of the public administration. Model 5 has significantly improved fit compared to Model 4 ($\chi^2(1) = 1146,39$, $p < 0,001$), with a decrease in total variance compared to Model 1 by 34,64 percent.

Analysing the residuals of Model 5 for protection of life and physical integrity, both the level 1 residuals and the level 2 residuals seem fairly normally distributed judging by Panel A of Figure 7. The level 1 residuals do seem to have a slight right skew. The autocorrelation plot of Panel B suggests that values within the same country remain strongly related over time. Predicted value-residual plots of the level 1 and level 2 residuals showed a fairly strong downward slope in the level 2 residuals.

Figure 7

QQ-plot and autocorrelation plot for protection of life and physical integrity.



Note: The standardised residuals are standardised with country-level standard deviations.

Table 5

Linear random intercept model with protection of life and physical integrity as dependent variable and year-fixed effects (N = 3406, n = 109, T = 1-55).

Variable	Model 3		Model 4		Model 5	
	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>
Intercept	45,08 (4,68)	<0,01	131,78 (45,18)	0,01	173,37 (38,53)	<0,01
Personalism	-17,48 (3,94)	<0,01	-16,96 (3,47)	<0,01	-9,33 (2,78)	<0,01
Rig. & Impart. Pub. Admin.					11,23 (1,53)	<0,01
Political Violence _{<i>t</i>-1}			-4,36 (1,09)	<0,01	-2,44 (1,00)	0,02
log ₁₀ (Population) _{<i>t</i>-1}			-15,08 (6,64)	0,04	-17,13 (5,60)	0,01
ln(GDP per cap) _{<i>t</i>-1}			1,27 (2,03)	0,54	-1,89 (1,73)	0,28
GDP Growth			-5,38 (2,51)	0,16	-4,02 (1,88)	0,11
Internal conflict (1 = conflict) _{<i>t</i>-1}			-1,59 (1,46)	0,26	-0,20 (1,25)	0,74
International conflict (1= conflict) _{<i>t</i>-1}			-1,00 (1,44)	0,29	-1,25 (1,69)	0,29
Regime type _{<i>t</i>-1}						
Personalist (ref)						
Military			-1,75 (4,06)	0,67	-2,45 (3,64)	0,57
Single-party			1,75 (5,07)	0,73	0,20 (3,50)	0,93
Monarchy			18,44 (5,53)	0,01	9,74 (3,60)	0,05
Random effect						
Between-country	423,04		333,09		316,61	
Within-country	137,09		119,74		84,75	
R_I^2 ^c	0,09		0,26		0,35	
ICC	0,76		0,74		0,79	
LR χ^2 Year-fixed ^b	248,25	<0,01	323,82	<0,01	320,39	<0,01
LR χ^2 ^d	243,69	<0,01	481,15	<0,01	1146,39	<0,01

Note: ^a Country-clustered standard errors. ^b Compares model with year effects to an equivalent model without them. ^c Proportional decrease in level 1 variance compared to Model 1.

^d Compares each model to the previous model.

Freedom of movement

The first two models are once again reported in the running text only. Model 1 with only random country intercepts for freedom of movement estimates within-country variance to be 0,20 while the estimated between-country variance is 1,23 ($ICC = 0,86$). The between-country mean does not differ significantly from zero ($b < 0,01$, $SE = 0,11$, 98,75% CI [-0,27; 0,27]). Model 2 introduces year-fixed effects, which contribute significantly to model fit ($\chi^2(54) = 382,52$, $p < 0,001$) and show that the mean level of freedom of movement is

significantly higher than it was in 1946 from 1990 onwards. Model 2 has a within-group variance of 0,18 and a between-group variance of 1,22, giving an intra-country correlation of 0,87 and decreasing total variance by 2,37 percent ($R_I^2 = 0,02$) compared to Model 1.

The first model reported in Table 6 is Model 3, which introduces the personalism variable. Countries with the highest level of personalism are estimated to have a level of freedom of movement that is 0,40 points lower ($SE = 0,12$, 98,75% CI [-0,70; -0,10]) than countries with the lowest possible level of personalism, given time-persistent differences between countries and across-country yearly changes in freedom of movement. This is a small-to-medium effect, taking into account the variation in freedom of movement and the fact that it compares the extremes of the personalism scale. The decrease in total variance compared to Model 1 is just 4,30 percent.

Model 4 is second in Table 6, and this model adds the lagged control variables. None of these control variables has a statistically significant relation to freedom of movement at $\alpha = 0,0125$. The control variables do seem to contribute to model fit together ($\chi^2(9) = 405,26$, $p < 0,001$), but the decrease in level 1 variance compared to Model 1 is just 5,65 percent. The estimated relation of personalism to freedom of movement rises somewhat ($b = -0,47$, $SE = 0,10$, 98,75% CI [-0,74; -0,20]) when controlling for the control variables in addition to time-persistent differences between countries and across-country developments in a year. The negative relation between personalism and freedom of movement fits my hypothesis.

Model 5 is the final model in Table 6, adding the rigour and impartiality of the public administration to test for mediation. The total variance of this model is quite a lot lower than that of Model 1 ($R_I^2 = 0,18$). Higher levels of rigour and impartiality of the public administration tend to be accompanied by greater freedom of movement ($b = 0,34$, $SE = 0,06$, 98,75% CI [0,19; 0,50]) controlling for personalism, the control variables, across-country changes by year and time-persistent differences between countries. This effect is rather large and larger than that of personalism. Controlled for the rigour and impartiality of the public administration as well as time-persistent differences between countries and across-country changes by year, the effect of personalism becomes smaller and statistically indistinguishable from zero ($b = -0,24$, $SE = 0,10$, 98,75% CI [-0,49; 0,01]). However, the confidence intervals for personalism in Model 4 and Model 5 still contain the point estimate of the other model, such that this change is not that meaningful.

For the residual diagnostics of freedom of movement, Panel A of Figure 8 shows that the level 1 residuals have a great deal of excess kurtosis. This means that the standard errors calculated for Model 5 (and probably the other models) are under-estimated. The level 2

residuals have a stretched lower tail, but resemble a normally distributed variable a lot more. The autocorrelation plot of Panel B suggests that values within the same country remain strongly related over time. Predicted value-residual plots of the level 1 and level 2 residuals showed a slight downward slope in the level 2 residuals.

Table 6

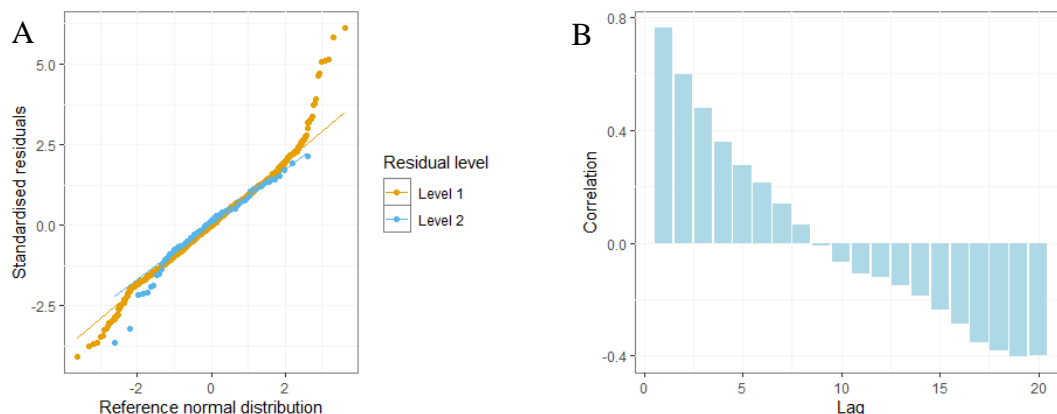
Linear random intercept model with freedom of movement as dependent variable and year-fixed effects (N = 3251, n = 107, T = 1-54).

Variable	Model 3		Model 4		Model 5	
	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>	<i>b</i> (SE) ^a	<i>p</i>
Intercept	0,06 (0,14)	0,68	1,00 (3,47)	0,77	2,37 (2,78)	0,40
Personalism	-0,40 (0,12)	<0,01	-0,47 (0,10)	<0,01	-0,24 (0,10)	0,01
Rig. & Impart. Pub. Admin.					0,34 (0,06)	<0,01
Political Violence			-0,10 (0,05)	0,03	-0,04 (0,04)	0,31
log ₁₀ (Population)			-0,33 (0,49)	0,50	-0,38 (0,39)	0,34
ln(GDP per cap)			0,19 (0,11)	0,08	0,10 (0,09)	0,27
GDP Growth			-0,11 (0,12)	0,40	-0,07 (0,09)	0,49
Internal conflict (1 = conflict)			-0,15 (0,07)	0,03	-0,11 (0,05)	0,05
International conflict (1= conflict)			-0,08 (0,04)	0,09	-0,08 (0,04)	0,06
Regime type						
Personalist (ref)						
Military			-0,26 (0,09)	0,01	-0,28 (0,09)	<0,01
Single-party			-0,37 (0,14)	0,01	-0,41 (0,12)	<0,01
Monarchy			0,24 (0,34)	0,49	-0,03 (0,19)	0,89
Random effect						
Between-country	1,20		1,20		1,05	
Within-country	0,18		0,16		0,12	
<i>R</i> _l ^{2c}	0,04		0,06		0,18	
<i>ICC</i>	0,87		0,88		0,89	
<i>LR</i> χ^2 Year-fixed ^b	433,70	<0,01	277,68	<0,01	261,13	<0,01
<i>LR</i> χ^2 ^d	99,89	<0,01	405,26	<0,01	762,32	<0,01

Note: ^a Country-clustered standard errors. ^b Compares model with year effects to an equivalent model without them. ^c Proportional decrease in level 1 variance compared to Model 1. ^d Compares each model to the previous model.

Figure 8

QQ-plot and autocorrelation plot for freedom of movement.



Note: The standardised residuals are standardised with country-level standard deviations.

Rigour and impartiality of the public administration

In the case of the rigour and impartiality of the public administration, Model 1 with only random country intercepts has a within-country variance of 0,33 and a between-country variance of 0,91 ($ICC = 0,73$). The between-country mean level of the rigour and impartiality of the public administration is -0,60 ($SE = 0,09$, 98,75% CI [-0,84; -0,37]). Model 2 adds year-fixed effects, but in this case doing so does little to improve model fit ($\chi^2(54) = 76,85$, $p = 0,022$). Indeed, total variance in Model 2 is just 0,14 percent lower than in Model 1.

Model 3 adds personalism, and this does improve model fit ($\chi^2(1) = 199,28$, $p < 0,001$) with the level 1 variance decreasing by 3,23 percent. The point estimate estimates the rigour and impartiality to be 0,75 lower ($SE = 0,15$, 98,75% CI [-1,14; -0,35]) in countries with the highest possible level of personalism as compared to the lowest possible level of personalism controlling for time-persistent differences between countries and across-country yearly changes. This is a small effect, and the estimate is rather uncertain. The year-fixed effects still do not significantly contribute to model fit ($\chi^2(54) = 70,79$, $p = 0,062$).

Model 4 controls the relation between personalism and the rigour and impartiality of the public administration for the same control variables used in the other models. Total variance decreases quite strongly ($R_I^2 = 0,16$). The estimated relation between personalism and the rigour and impartiality of the public administration controlling for time-persistent country differences and across-country year effects as well as these control variables is somewhat smaller than that in Model 3 ($b = -0,67$, $SE = 0,15$, 98,75% CI [-1,06; -0,28]), but due to the uncertainty of the estimate this change does not seem overly meaningful. Higher levels of political violence in the previous year are associated with lower levels of rigour and

impartiality of the public administration in this year ($b = -0,17$, $SE = 0,06$, 98,75% CI [-0,32; -0,02]) given the level of personalism, the other control variables, time-persistent differences between countries and across-country differences over years, but once again the size of the effect is hard to determine. In this model the year-fixed effects are statistically significant ($\chi^2(54) = 109,65$, $p < 0,001$), but since they only first become so in this model it is questionable how much weight should be given to that.

The level 1 residuals of Model 5 of the rigour and impartiality of the public administration seem to be right-skewed, shown in Panel A of Figure 9. Although less severe, such a right-skew also seems to appear in the level 2 residuals. Panel B of Figure 9 shows the autocorrelation plot of level 1 residuals, and this suggests that residuals are still very strongly related to each other over time within a country. Predicted value-residual plots of the level 1 and level 2 residuals showed a slight downward slope in the level 2 residuals.

Table 7

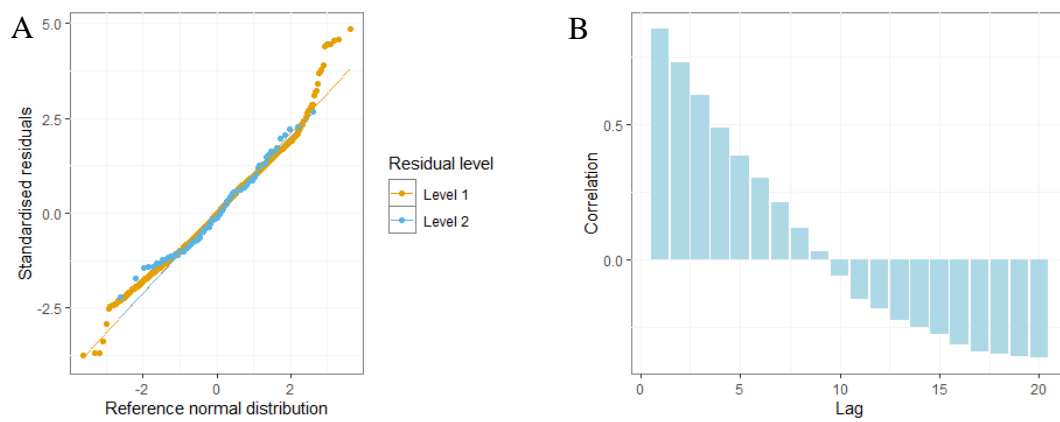
Linear random intercept model with rigour and impartiality of the public administration as dependent variable and year-fixed effects (N = 3406, n = 109, T = 1-55).

Variable	Model 3		Model 4	
	b (SE) ^a	p	b (SE) ^a	p
Intercept	-0,46 (0,16)	0,15	-2,28 (2,49)	0,36
Personalism	-0,71 (0,15)	<0,01	-0,67 (0,15)	<0,01
Political Violence			-0,17 (0,06)	0,01
log ₁₀ (Population)			-0,02 (0,33)	0,95
ln(GDP per cap)			0,27 (0,13)	0,06
GDP Growth			-0,12 (0,16)	0,45
Internal conflict (1 = conflict)			-0,13 (0,11)	0,27
International conflict (1= conflict)			0,02 (0,08)	0,78
Regime type				
Personalist (ref)				
Military			0,06 (0,15)	0,69
Single-party			0,14 (0,22)	0,54
Monarchy			0,77 (0,48)	0,15
Random effect				
Between-country	0,89		0,76	
Within-country	0,31		0,27	
R_l^2 ^c	0,03		0,16	
ICC	0,74		0,74	
LR χ^2 Year-fixed ^b	70,79	0,06	109,65	<0,01
LR χ^2 ^d	199,28	<0,01	429,39	<0,01

Note: ^a Country-clustered standard errors. ^b Compares model with year effects to an equivalent model without them. ^c Proportional decrease in level 1 variance compared to Model 1. ^d Compares each model to the previous model.

Figure 9

QQ-plot and autocorrelation plot for rigour and impartiality of the public administration.



Note: The standardised residuals are standardised with country-level standard deviations.

Conclusion & Discussion:

Based on the results of the random intercept regressions, higher levels of personalism in a country tend to go along with somewhat lower freedom of expression, freedom of assembly, protection of life and physical integrity and freedom of movement. It might be that this is partly because higher levels of personalism in a country are associated to lesser competence of coercive institutions, but my evidence for this remains rather inconclusive. I need to reject my hypothesis of a curvilinear association between personalism and freedom of expression, but the extent to which the theory underlying the other hypotheses is corroborated is also limited. This means that my results are in line with the findings of Frantz et al. (2019) and broaden them somewhat by looking at civil liberties instead of repression.

I am not able to come to all too firm conclusions with regards to the role of the competence of coercive institutions in explaining reduced civil liberties in more personalised regimes. However, that my results at least suggest an effect does give me access to evidence for a mechanism which Frantz et al. (2019) do not give. This deepens their results, and in order to understand why personalism is associated to worse conditions for the inhabitants of a country it is important to continue looking into evidence for possible mechanisms. The result that my operationalisation of the competence of coercive institutions, the rigour and impartiality of the public administration, is related to greater civil liberties is itself in line with the predictions of, among others, Greitens (2016).

A conceptually important issue is that the operationalisation of personalism I use (Geddes et al., 2018; Wright, 2021) leans heavily on concrete actions connected with personalisation, such as the creation of a support party or purges of the military. Frantz et al. (2019) suggest that this operationalisation is at least not intrinsically connected to repression by including indicators of repression as indicators of personalism. Since I operationalise the competence of coercive institutions with a measure of rigour and impartiality of the public administration, which measures the respect of public officials for the law and unbiased administration of it (Coppedge et al., 2021a, pp. 175-176), this measure should also be distinct from personalism. The personalism measure is based on the relation between the dictator and his government, while the rigour and impartiality of the public administration operationalises the relation between the government and the general populace.

However, going beyond definitions and towards my theoretical argument, if a dictator undermines the competence of coercive institutions to allow further personalisation this suggests that a dictator that does not undermine the competence of coercive institutions might

not be able to personalise rule as much. This is not a necessary relation, but it muddles the direction of causality and means that personalism and the competence of coercive institutions might be hard to disentangle empirically. This might explain why I did not only find (weak) evidence for the mechanism where I expected to do so, but also where I did not expect it.

A conundrum in my research setup is that I test for the effect of personalism on four civil liberties that are theoretically and empirically distinct, but are sufficiently related that the effect of personalism on one might bias the estimated effect of personalism on another. At the same time, adding the other civil liberties to the model of any one of them would create the rather questionable theoretical setup of looking at the effect of personalism on a civil liberty in those situations where the other civil liberties remain unaffected by it. Since the effects of personalism are fairly similar across the civil liberties, while the relations between civil liberties are strong but not extremely so, the extreme case of some civil liberties only being related to personalism through their relation to other civil liberties seems unlikely. However, this does not solve the underlying theoretical problem. A strategy to address this problem in future is to use multivariate multilevel models (Snijders & Bosker, 2012, pp. 282-288), which explicitly incorporate the possibility of multiple related outcomes.

Another problem in my model specification is that I have not modelled the dependence of values within a country from one year to the next. The level 1 residuals still showed strong dependence over time. This means that the amount of information my data contains is far lower than my model assumes, such that my statistical tests give statistically significant results more often than is justified. Of course, this can very plausibly mean that my conclusion that personalism affects civil liberties through the competence of coercive institutions is not statistically justifiable.

Beyond implausibly assuming within-country independence over time, another statistical limitation is the fact that I have had to assume that observations in different countries are independent from each other. While such independence is not problematic for individual humans randomly sampled from a large population, for a dataset containing all autocracies for the period under study this seems rather less plausible. This is exacerbated by the fact that countries might not only be similar because they are affected by a shared environment (e.g., interference from democratic states), but autocratic countries also have an interest to interfere in the domestic affairs of each other. I have assumed that personalism and civil liberties in the various members of the Warsaw pact developed independently, for example, while the Berlin Uprising of 1953, the Prague Spring of 1968, and the revolutions of 1989 make this assumption seem rather implausible.

An interesting extension to test the mechanisms I have presented here is to operationalise the threat of an elite coup or the extent to which the dictator is able to monitor their elite. I have only been able to study the proposed effect of elite coup threat in a very indirect way that requires a lot of my assumptions to be correct. I could not empirically distinguish the ability of a dictator to monitor their elite at all. Being able to look into the role of these mechanisms in more depth would certainly enrich a study of the consequences of personalism.

That the effect that I find is rather small and very uncertain makes clear practical implications hard to draw. Assuming that the consequences for human well-being nonetheless recommend some type of intervention I can make a few suggestions. If personalism in autocratic countries is a threat to civil liberties, it seems wise for liberal democracies to promote a more dispersed power structure in autocratic countries instead of treating all forms of autocratic government as equally undesirable. Given that the competence of coercive institutions might be one mechanism through which personalism leads to less respect for civil liberties, a method for achieving the first goal that can reduce suffering in autocracies in and of itself might be to help and pressure autocracies to establish a professional government structure that is not dominated by clientelist practices and can exert soft power. However, for both these approaches the threat of perverse consequences is considerable. Keeping power from getting overly concentrated can itself disadvantage the population of a countries if it leads to destabilisation of a regime and possibly civil war. And a more professional government apparatus is only beneficial to civil liberties if it uses its capacities for more restrained methods of control, instead of more rigorous ones. And insofar as professionalising government conflicts with personalisation, getting a dictator to cooperate will be very difficult.

To summarise, how much power lies in the hands of a dictator seems to have some deleterious consequences for freedoms of expression, assembly, movement and the protection of life and physical integrity. This seems to happen partly because dictators suppress the competence of their coercive institutions, but which other mechanisms are at play remains to be seen.

References:

- Acemoglu, D., Egorov, G. & Sonin, K. (2008). Coalition formation in non-democracies. *Review of Economic Studies*, 75(4), 987-1009. <https://doi.org/10.1111/j.1467-937X.2008.00503.x>
- Acemoglu, D., Egorov, G., & Sonin, K. (2009). Political economy under weak institutions: Do juntas lead to personal rule? *American Economic Review*, 99(2), 298-303. <https://doi.org/10.1257/aer.99.2.298>
- Acemoglu, D., Egorov, G., & Sonin, K. (2012). Dynamics and stability of constitutions, coalitions, and clubs. *American Economic Review*, 102(4), 1446-1476. <https://doi.org/10.1257/aer.102.4.1446>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Baturo, A., & Elkink, J. A. (2016). Dynamics of regime personalisation and patron-client networks in Russia, 1999–2014. *Post-Soviet Affairs*, 32(1), 75-98. <https://doi.org/10.1080/1060586x.2015.1032532>
- Boleslavsky, R., Shadmehr, M., & Sonin, K. (2021). Media freedom in the shadow of a coup. *Journal of the European Economic Association*, 19(3), 1782-1815. <https://doi.org/10.1093/jeea/jvaa040>
- Bolt, J., Inklaar, R., de Jong, H., & van Zanden, J. L. (2018). *Rebasing 'Maddison': New income comparisons and the shape of long-run economic development* (GGDC Research Memorandum 174). Groningen Growth and Development Centre. https://www.rug.nl/ggdc/html_publications/memorandum/gd174.pdf

Chin, J. (2020, May 8). *Five myths about coups: No, they're not found only in poorer nations.*

The Washington Post. https://www.washingtonpost.com/outlook/five-myths/five-myths-about-coups/2020/05/07/9c64ee04-8f1d-11ea-B9c0-73b93422d691_story.html

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Lührmann, A., Maerz, S. F., ... Ziblatt, D. (2021a). *V-Dem Codebook* [Code Book] (version 11.1). Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Lührmann, A., Maerz, S. F., ... Ziblatt, D. (2021b). *V-Dem Country–Year Dataset* [Data set] (version 11.1). Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds21>

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Lührmann, A., Maerz, S. F., ... Ziblatt, D. (2021c). *V-Dem Methodology* (version 11). Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds21>

Davenport, C. (2007a). State repression and political order. *Annual Review of Political Science*, 10, 1-23. <https://doi.org/10.1146/annurev.polisci.10.101405.143216>

Davenport, C. (2007b). State repression and the tyrannical peace. *Journal of Peace Research*, 44(4), 485-504. <https://doi.org/10.1177/0022343307078940>

Davenport, C., RezaeeDaryakenari, B., & Wood, R. M. (2021). Tenure through tyranny? Repression, dissent, and leader removal in Africa and Latin America, 1990-2006.

Journal of Global Security Studies, 7(1), Article ogab023.

<https://doi.org/10.1093/jogss/ogab023>

Egorov, G., & Sonin, K. (2020). *The political economics of non-democracy* (Working Paper 27949). National Bureau of Economic Research. <http://www.nber.org/papers/w27949>

Frantz, E., Kendall-Taylor, A., Wright, J., & Xu, X. (2019). Personalisation of power and repression in dictatorships. *The Journal of Politics*, 82(1), 372-377.

<https://doi.org/10.1086/706049>

Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic breakdown and regime transitions. *Perspectives on Politics*, 12(2), 313-331.

<https://doi.org/10.1017/S1537592714000851>

Geddes, B., Wright, J., & Franz, E. (2018). *How dictatorships work: Power, personalisation and collapse*. Cambridge University Press.

<https://doi.org/10.1017/9781316336182.004>

Gehlbach, S., Sonin, K., & Svulik, M. W. (2016). Formal models of nondemocratic politics. *Annual Review of Political Science*, 19, 565-584. [https://doi.org/10.1146/annurev-](https://doi.org/10.1146/annurev-polisci-042114-014927)

[polisci-042114-014927](https://doi.org/10.1146/annurev-polisci-042114-014927)

Greitens, S. C. (2016). *Dictators and their secret police: Coercive institutions and state violence*. Cambridge University Press.

<https://doi.org/10.1017/CBO9781316489031.002>

Hill, D. W., & Jones, Z. M. (2014). An empirical evaluation of explanations for state repression. *The American Political Science Review*, 108(3), 661-687.

<https://doi.org/10.1017/S000305541>

Hollyer, J. R., Rosendorff, B. P., & Vreeland, J. R. (2019). Why do autocrats disclose?

Economic transparency and inter-elite politics in the shadow of mass unrest. *Journal of Conflict Resolution*, 63(6), 1488-1516. <https://doi.org/10.1177/0022002718792602>

- Jandoc, K., & Juarez, R. (2019). An experimental study of self-enforcing coalitions. *Games*, 10(3), Article 31. <https://doi.org/10.3390/g10030031>
- Levitsky, S., & Way, L. A. (2016). Durable authoritarianism. I. O. Fioretos, T. G. Faletti, & A. Sheingate (Eds.), *The Oxford handbook of historical institutionalism*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199662814.013.12>
- Libman, A. (2020). Temporality of authoritarian regimes. In K. H. Goetz (Ed.), *The Oxford handbook of time and politics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190862084.013.29>
- Newson, A., & Trebbi, F. (2018). Authoritarian elites. *Canadian Journal of Economics*, 51(4), 1088-1117. <https://doi.org/10.1111/caje.12362>
- MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3), 305-325. [https://doi.org/10.1016/0304-4076\(85\)90158-7](https://doi.org/10.1016/0304-4076(85)90158-7)
- McDermott, K. (2014). Stalin and Stalinism. In S. A. Smith (Ed.), *The Oxford handbook of the history of communism* (pp. 72-87). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199602056.013.007>
- Møller, J., & Skaaning, S.-E. (2013a). Autocracies, democracies and the violation of civil liberties. *Democratization*, 20(1), 82-106. <https://doi.org/10.1080/13510347.2013.738863>
- Møller, J., & Skaaning, S.-E. (2013b). Respect for civil liberties during the third wave of democratization: Presenting a new dataset. *Social Indicators Research*, 117(3), 1069-1087. <https://doi.org/10.1007/s11205-013-0391-y>
- Møller, J., & Skaaning, S.-E. (2013c). Single-party autocracies, ideology, and repression. *Statsvetenskaplig Tidskrift*, 115(4), 345-363. <https://journals.lub.lu.se/st/article/view/9742>

- Newson, A., & Trebbi, F. (2018). Authoritarian elites. *Canadian Journal of Economics*, 51(4), 1088-1117. <https://doi.org/10.1111/caje.12362>
- Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y.-t., Medzihorsky, J., Krusell, J., Miri, F., & van Römer, J. (2021). *The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (Working Paper No. 21, 6th ed.). Varieties of Democracy (V-Dem) Project.
https://web.archive.org/web/20211021062126/https://www.v-dem.net/media/filer_public/d4/d2/d4d27532-96a7-4f8c-8028-4c356fb1b2ca/wp_21_6th_edition.pdf
- Pustejovsky, J. E., & Tipton, E. (2016). Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics*, 36(4), 672-683. <https://doi.org/10.1080/07350015.2016.1247004>
- R Core Team. (2021). *R: A language and environment for statistical computing*. (Version 4.1.1) [Computer software]. R Foundation for Statistical Computing <https://www.R-project.org/>
- Sinkkonen, E. (2021). Dynamic dictators: improving the research agenda on autocratization and authoritarian resilience. *Democratization*, 28(6), 1172-1190.
<https://doi.org/10.1080/13510347.2021.1903881>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Svolik, M. W. (2012). *The politics of authoritarian rule*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139176040.003>
- Turner, T. & Young, C. (1985). *The rise and decline of the Zairian State*. University of Wisconsin Press.

Universal Declaration of Human Rights, December 10, 1948,

<https://www.un.org/sites/un2.un.org/files/udhr.pdf>

Wright, J. (2021). The latent characteristics that structure autocratic rule. *Political Science*

Research and Methods, 9(1), 1-19. <https://doi.org/10.1017/psrm.2019.50>

Zu, Z. (2020). *Logic of selective repression: How do elite purges affect state violence in*

authoritarian regimes [Master's Thesis, Duke University]. DukeSpace.

[https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/20790/Zu_duke_0066
N_15620.pdf?sequence=1&isAllowed=y](https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/20790/Zu_duke_0066_N_15620.pdf?sequence=1&isAllowed=y)

Appendix: R script

This appendix contains the R script used to run the analyses for this thesis. The script starts at the next page. The datasets used by the script are “GWF+personalism-scores.csv” which can be retrieved from <https://tinyurl.com/277vebkf>, “GWF Autocratic Regimes.xlsx”, which can be retrieved from <https://tinyurl.com/2cpwmwvt>, and “GWFtscs.txt”, which can be retrieved from <https://tinyurl.com/23krvz7a>.

The shortened URLs above are more useful for manual entry. In case they do not work or the full URL is preferred, these are:

V-Dem: www.v-dem.net/media/datasets/Country_Year_V-Dem_Full_others_R_v11.1.zip

Geddes et al. (2018): www.sites.psu.edu/dictators/wp-content/uploads/sites/12570/2016/05/GWF-Autocratic-Regimes-1.2.zip

Wright (2021): www.sites.psu.edu/wright/files/2019/11/GWF-time-vary-personalism.zip

```

1 # Author: Siebren Kooistra
2 # Date: 05-06-2022
3 # Goal: Construct a dataset on personalism and civil liberties using
4 # Geddes et al. (2014, 2018) and V-Dem (Coppedge et al., 2021) and
5 # carry out analyses on the associations between personalism and civil
6 # liberties
7
8 # Preparations: clean workspace and mount libraries.
9 rm(list = ls())
10 graphics.off()
11 gc()
12 library(DescTools)
13 library(readxl)
14 library(tidyverse)
15 library(GGally)
16 library(lmtest)
17 library(ggrepel)
18 library(mice)
19 library(naniar)
20 library(lme4)
21 library(arsenal)
22 library(plm)
23 library(clubSandwich)
24
25 # Create visual elements for later use
26 blue_light <- theme(panel.background = element_rect(fill = "white",
27                                                    colour = "grey50"),
28                   panel.grid.major = element_line(colour = "whitesmoke"),
29                   panel.grid.minor = element_line(colour = "whitesmoke"),
30                   legend.key = element_rect(fill = "white",
31                                             colour = "grey50"))
32 lines_palette <- c("darkgreen", "blue", "lightblue", "lightblue", "red")
33
34 #####
35 ##### DATASET ASSEMBLY AND DESCRIPTION #####
36 #####
37
38 ##### Dataset assembly #####
39
40 # First dataset: GWF with personalism
41
42 # Load GWF autocracies dataset with personalism data and specify
43 # which variables to use
44 GWF_personalism <- read.csv("GWF+personalism-scores.csv")
45 GWF_personalism_variables_of_interest <- c("cowcode", "gwf_country", "year",
46                                           "dataID", "latent_personalism",
47                                           "paramil_pers", "sectyapp_pers",
48                                           "officepers", "partyrbrstmp",
49                                           "militparty_newparty", "milnotrial",
50                                           "milmerit_pers", "partyexcom_pers")
51
52 # Define a new variable dataID to match V-Dem country-years to GWF
53 # country-years. In this case, V-Dem takes the separation of Yemen into two
54 # countries from 1918 to 1989 into account by giving the unified Yemen
55 # existing until 1918 and from 1990 a COW code of 679. Since this is more
56 # informative than the constant GWF code of 678, the GWF code is overwritten
57 # as 679 from 1990 to 2010.
58 GWF_personalism$cowcode[GWF_personalism$cowcode == 678 &
59                       GWF_personalism$year %in% 1990:2010] <- 679

```

```

60
61 # Second dataset: GWF regime types
62
63 # Load the GWF regime type dataset and specify which variables to use
64 GWF_regime_type <- read.delim("GWFtscs.txt")
65 GWF_regime_type_variables_of_interest <- c("dataID", "gwf_party",
66                                           "gwf_personal", "gwf_military",
67                                           "gwf_monarch")
68
69 # Apply the same correction for South Yemen as used in the personalism
70 # dataset to the regime type dataset
71 GWF_regime_type$cowcode[GWF_regime_type$cowcode == 678 &
72                          GWF_regime_type$year %in% 1990:2010] <- 679
73
74 # Third dataset: V-Dem
75
76 # Load V-Dem datafile and specify which variables to use
77 VDem <- readRDS("V-Dem-CY-Full+Others-v11.1.rds")
78 VDem_variables_of_interest <- c("COWcode", "dataID", "v2x_clphy", "v2cltort",
79                                "v2clkill", "v2caassemb", "v2clfmmove",
80                                "v2cldmovem", "v2cldmovew", "v2x_freexp",
81                                "v2mecenefm", "v2meharjrn", "v2meslfcen",
82                                "v2xcl_disc", "v2clacfree", "e_mipopula",
83                                "e_miinteco", "e_miinterc", "e_migdppc1n",
84                                "e_migdpgro", "v2caviol", "v2clrspct",
85                                "v2cldiscm", "v2cldiscw", "lag_e_miinteco",
86                                "lag_e_miinterc", "lag_e_migdppc1n",
87                                "lag_e_migdpgro", "lag_v2caviol",
88                                "lag_e_mipopula")
89
90 # Create variables to match the three datasets
91 GWF_regime_type$dataID <- paste(GWF_regime_type$cowcode, GWF_regime_type$year)
92 GWF_personalism$dataID <- paste(GWF_personalism$cowcode, GWF_personalism$year)
93 VDem$dataID <- paste(VDem$COWcode, VDem$year)
94
95 # Lag variables
96 VDempdata <- pdata.frame(VDem, index = c("COWcode", "year"))
97 VDem$lag_e_miinteco <- plm::lag(VDem$e_miinteco)
98 VDem$lag_e_miinterc <- plm::lag(VDem$e_miinterc)
99 VDem$lag_e_migdppc1n <- plm::lag(VDem$e_migdppc1n)
100 VDem$lag_e_migdpgro <- plm::lag(VDem$e_migdpgro)
101 VDem$lag_v2caviol <- plm::lag(VDem$v2caviol)
102 VDem$lag_e_mipopula <- plm::lag(VDem$e_mipopula)
103
104 # Combine the GWF data and the V-Dem data
105
106 # Subset the part of the V-Dem dataset that is relevant for the GWF data
107 VDem_Repression <- VDem[VDem$dataID %in% GWF_personalism$dataID,
108                          c("historical_date",
109                              VDem_variables_of_interest)]
109
110 # Combine the datasets
111 autocracy_data <- cbind(GWF_personalism[order(GWF_personalism$dataID),
112                                           GWF_personalism_variables_of_interest],
113                          GWF_regime_type[order(GWF_regime_type$dataID),
114                                           GWF_regime_type_variables_of_interest],
115                          VDem_Repression[order(VDem_Repression$dataID),
116                                           VDem_variables_of_interest])
117
118

```

```

119 # Check for sorting errors.
120 TEST <- all(autocracy_data$cowcode == autocracy_data$COWcode)
121
122 # Variable operations
123 # Create a log-10 population variable
124 autocracy_data$log10pop <- log10(autocracy_data$e_mipopula * 1000)
125 autocracy_data$lag_log10pop <- log10(autocracy_data$lag_e_mipopula * 1000)
126
127 # Compute the scale variable for freedom of movement
128 autocracy_data$freedom_movement <-
129   rowMeans(autocracy_data[, c("v2c1fmove", "v2c1dmovem", "v2c1dmovew")])
130
131 # Multiply freedom of expression and protection of life and physical integrity
132 # scales by 100 to avoid errors due to floating point arithmetic
133 autocracy_data$free_expr_x100 <- autocracy_data$v2x_freexp * 100
134 autocracy_data$life_phys_x100 <- autocracy_data$v2x_clphy * 100
135
136 # Tidy up dataset
137 autocracy_data <- autocracy_data[, colnames(autocracy_data) %nin% c("dataID",
138   "COWcode")]
139
140 # Subset complete cases
141 datacomplete <- autocracy_data[complete.cases(autocracy_data),]
142
143 # Write dataset into csv file
144 write.csv(autocracy_data,
145   "autocracies_personalism_civilrights_data_S_Kooistra.csv")
146 write.csv(datacomplete,
147   "complete_cases_autocracies_personalism_civilrights_S_Kooistra.csv")
148
149 # Clean up to save memory.
150 rm(GWF_personalism, GWF_personalism_variables_of_interest,
151   GWF_regime_type, GWF_regime_type_variables_of_interest,
152   VDem, VDem_Repression, VDem_variables_of_interest,
153   datacomplete, autocracy_data) # Also remove dataframes to re-load and
154   # remove pdata.frame properties
155 gc()
156
157 # Re-load dataframes
158 autocracy_data <-
159   read.csv("autocracies_personalism_civilrights_data_S_Kooistra.csv")
160 datacomplete <-
161   read.csv("complete_cases_autocracies_personalism_civilrights_S_Kooistra.csv")
162 ##### Personalism descriptives #####
163
164 # Item description and scale evaluation
165
166 # Military promotions
167
168 # Table and chi-square test
169 milmerit_pers_table <- table(autocracy_data$milmerit_pers, autocracy_data$year)
170 milmerit_pers_table
171 summary(milmerit_pers_table)
172
173 # Stacked barplot per year
174 ggplot(autocracy_data, aes(x = year, fill = as.factor(milmerit_pers))) +
175   geom_bar(position = "fill") + blue_light +
176   labs(x = "Year", y = "Proportion", fill = "Military promotions") +
177   scale_fill_manual(values = c("tomato", "lightblue", "darkgreen"),

```

```

178   labels = c("No forced retirement, \nno loyalty-based promotion, \nor no
179   military",
180             "Loyalty-based promotion",
181             "In-group promotion \nor forced retirement"))
182
183   # Recode military promotion variable to dummy
184   autocracy_data$milmerit_pers_twocat <- recode(autocracy_data$milmerit_pers,
185                                               `0` = 0, `1` = 0, `2` = 1)
186
187   # Investigate distribution of dummy
188   milmerit_pers_twocat_table <- table(autocracy_data$milmerit_pers_twocat,
189                                     autocracy_data$year)
190   milmerit_pers_twocat_table
191   summary(milmerit_pers_twocat_table)
192
193   # Stacked barplot per year for dummy
194   ggplot(autocracy_data, aes(x = year, fill = as.factor(milmerit_pers_twocat))) +
195     geom_bar(position = "fill") + blue_light +
196     labs(x = "Year", y = "Proportion", fill = "Military promotions") +
197     scale_fill_manual(values = c("lightblue", "darkgreen"),
198                       labels = c("No forced retirement, \nno loyalty-based
199   promotion, \nor no military, \nor loyalty-based promotion",
200     "In-group promotion \nor forced retirement"))
201
202   # Military purges
203
204   # Table and chi-square test
205   milnotrial_table <- table(autocracy_data$milnotrial, autocracy_data$year)
206   milnotrial_table
207   summary(milnotrial_table)
208   chisq.test(autocracy_data$milnotrial, autocracy_data$year)$stdres
209
210   # Stacked barplot per year
211   ggplot(autocracy_data, aes(x = year, fill = as.factor(milnotrial))) +
212     geom_bar(position = "fill") + blue_light +
213     labs(x = "Year", y = "Proportion", fill = "Military purges") +
214     scale_fill_manual(values = c("lightblue", "darkgreen"),
215                       labels = c("No imprisonment or killing, \nno military, \nor
216   foreign military",
217     "Officer imprisonment and/or killing"))
218
219   # Support party
220
221   # Table and chi-square test
222   militparty_newparty_table <- table(autocracy_data$militparty_newparty,
223                                     autocracy_data$year)
224   militparty_newparty_table
225   summary(militparty_newparty_table)
226   chisq.test(autocracy_data$militparty_newparty, autocracy_data$year)$stdres
227   chisq.test(autocracy_data$militparty_newparty, autocracy_data$year)$expected
228   fisher.test(autocracy_data$militparty_newparty, autocracy_data$year,
229              simulate.p.value = TRUE, B = 10000)
230
231   # Stacked barplot per year
232   ggplot(autocracy_data, aes(x = year, fill = as.factor(militparty_newparty))) +
233     geom_bar(position = "fill") + blue_light +
234     labs(x = "Year", y = "Proportion", fill = "Party creation") +
235     scale_fill_manual(values = c("lightblue", "darkgreen"),
236                       labels = c("No support party created",

```

```

237         "Support party created"))
238
239 #     Executive committee appointments
240
241 #     Table and chi-square test
242 partyexcom_pers_table <- table(autocracy_data$partyexcom_pers,
243                               autocracy_data$year)
244 partyexcom_pers_table
245 summary(partyexcom_pers_table)
246 chisq.test(autocracy_data$partyexcom_pers, autocracy_data$year)$stdres
247
248 #     Stacked barplot per year
249 ggplot(autocracy_data, aes(x = year, fill = as.factor(partyexcom_pers))) +
250   geom_bar(position = "fill") + blue_light +
251   labs(x = "Year", y = "Proportion",
252        fill = "Party executive committee control") +
253   scale_fill_manual(values = c("lightblue", "darkgreen"),
254                    labels = c("Regime leader does not choose \nparty executive
255 committee",
256                               "Regime leader chooses \nparty executive
257 committee"))
258
259 #     Executive committee rubberstamp
260
261 #     Check coding
262 cor(autocracy_data$partyrbstmp,
263     autocracy_data[, c("latent_personalism", "paramil_pers",
264                       "sectyapp_pers", "officepers", "partyrbstmp",
265                       "militparty_newparty", "milnotrial",
266                       "milmerit_pers", "partyexcom_pers")])
267
268 #     Table and chi-square test
269 partyrbstmp_table <- table(autocracy_data$partyrbstmp,
270                            autocracy_data$year)
271 partyrbstmp_table
272 summary(partyrbstmp_table)
273
274 #     Stacked barplot per year
275 ggplot(autocracy_data, aes(x = year, fill = as.factor(partyrbstmp))) +
276   geom_bar(position = "fill") + blue_light +
277   labs(x = "Year", y = "Proportion",
278        fill = "Party executive committee independence") +
279   scale_fill_manual(values = c("lightblue", "darkgreen"),
280                    labels = c("The party executive committee \nhas some policy
281 independence \nfrom the regime leader",
282                               "The party executive committee \nis a 'rubber
283 stamp' \nor does not exist"))
284
285 #     Discretion over high office appointments
286
287 #     Table and chi-square test
288 officepers_table <- table(autocracy_data$officepers,
289                           autocracy_data$year)
290 officepers_table
291 summary(officepers_table)
292 chisq.test(autocracy_data$officepers, autocracy_data$year)$stdres
293
294 #     Stacked barplot per year
295 ggplot(autocracy_data, aes(x = year, fill = as.factor(officepers))) +

```

```

296   geom_bar(position = "fill") + blue_light +
297   labs(x = "Year", y = "Proportion",
298        fill = "Regime leader discretion over high office appointments") +
299   scale_fill_manual(values = c("lightblue", "darkgreen"),
300                     labels = c("Regime leader does not have discretion \nover
301 appointments to high office",
302                                "Regime leader has discretion over \nhigh office
303 appointments or appoints \nrelatives to these positions"))
304
305 #   Personalised control over the security apparatus
306
307 #   Table and chi-square test
308 sectyapp_pers_table <- table(autocracy_data$sectyapp_pers,
309                             autocracy_data$year)
310 sectyapp_pers_table
311 summary(sectyapp_pers_table)
312 chisq.test(autocracy_data$sectyapp_pers, autocracy_data$year)$stdres
313
314 #   Stacked barplot per year
315 ggplot(autocracy_data, aes(x = year, fill = as.factor(sectyapp_pers))) +
316   geom_bar(position = "fill") + blue_light +
317   labs(x = "Year", y = "Proportion",
318        fill = "Regime leader's personal control over the security apparatus") +
319   scale_fill_manual(values = c("lightblue", "darkgreen"),
320                     labels = c("Security apparatus is not controlled \npersonally
321 by the regime leader",
322                                "Security apparatus is controlled \npersonally by
323 the regime leader"))
324
325 #   Loyal paramilitary forces
326
327 #   Table and chi-square test
328 paramil_pers_table <- table(autocracy_data$paramil_pers,
329                             autocracy_data$year)
330 paramil_pers_table
331 summary(paramil_pers_table)
332 chisq.test(autocracy_data$paramil_pers, autocracy_data$year)$stdres
333
334 #   Stacked barplot per year
335 ggplot(autocracy_data, aes(x = year, fill = as.factor(paramil_pers))) +
336   geom_bar(position = "fill") + blue_light +
337   labs(x = "Year", y = "Proportion",
338        fill = "Creation of paramilitary forces loyal to the regime leader") +
339   scale_fill_manual(values = c("lightblue", "darkgreen"),
340                     labels = c("Regime leader does not create \nparamilitary
341 forces, \na president's guard, \nor new security forces \napparently loyal to
342 himself",
343                                "Regime leader creates \nparamilitary forces, \na
344 president's guard, \nor new security forces \napparently loyal to himself"))
345
346 #   Reliability analysis
347
348 #   Calculate Cronbach's Alpha by year
349 personalism_cronbachalphas <-
350   by(data = autocracy_data[, c("milmerit_pers_twocat", "milnotrial",
351                               "militparty_newparty", "partyexcom_pers",
352                               "partyrbrstmp", "officepers",
353                               "sectyapp_pers", "paramil_pers")],
354       INDICES = autocracy_data$year,

```



```

355     FUN = CronbachAlpha, cond = TRUE)
356
357 #       Create table for calculated Cronbach's Alpha
358 table_personalism_cronbachalphas <-
359   tibble(year = sort(unique(datacomplete$year)),
360         "All items" = NA,
361         "Military promotion strategies " = NA,
362         "Military purges" = NA,
363         "Support party creation" = NA,
364         "Party executive committee control" = NA,
365         "Party executive committee rubberstamp" = NA,
366         "Regime leader discretion over high office appointments" = NA,
367         "Personalised control over security apparatus" = NA,
368         "Loyal paramilitary forces" = NA
369   )
370
371 #       Fill table
372 for (y in sort(unique(autocracy_data$year))) {
373   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
374   "All items"] <-
375     personalism_cronbachalphas[[as.character(y)]]["unconditional"]
376   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
377   "Military promotion strategies"] <-
378     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
379   ]][1, "Cronbach Alpha"]
380   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
381   "Military purges"] <-
382     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
383   ]][2, "Cronbach Alpha"]
384   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
385   "Support party creation"] <-
386     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
387   ]][3, "Cronbach Alpha"]
388   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
389   "Party executive committee control"] <-
390     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
391   ]][4, "Cronbach Alpha"]
392   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
393   "Party executive committee rubberstamp"] <-
394     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
395   ]][5, "Cronbach Alpha"]
396   table_personalism_cronbachalphas[
397     table_personalism_cronbachalphas$year == y,
398     "Regime leader discretion over high office appointments"] <-
399     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
400   ]][6, "Cronbach Alpha"]
401   table_personalism_cronbachalphas[
402     table_personalism_cronbachalphas$year == y,
403     "Personalised control over security apparatus"] <-
404     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
405   ]][7, "Cronbach Alpha"]
406   table_personalism_cronbachalphas[table_personalism_cronbachalphas$year == y,
407   "Loyal paramilitary forces"] <-
408     personalism_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
409   ]][8, "Cronbach Alpha"]
410 }
411
412 #       Pivot table for use in plotting
413 longtab_personalism_cronbachalphas <-

```

```

414     pivot_longer(table_personalism_cronbachalphas, 2:10, names_to = "Type",
415                   values_to = "Value")
416
417 #     Plot Cronbach's Alpha by year
418 ggplot(longtab_personalism_cronbachalphas, aes(x = year, y = Value,
419                                               colour = Type)) +
420     geom_line() + blue_light
421
422 #     Calculate range of yearly Cronbach's Alphas
423 range(table_personalism_cronbachalphas$"All items" )
424
425 #     Scale descriptives
426
427 #     Entire dataset
428
429 #     Draw time series for Argentina, China, Congo/Zaire, Persia/Iran, Portugal
430 #     and the Soviet Union and Russia
431 ggplot(autocracy_data[autocracy_data$cowcode %in%
432                   c(160, 710, 490, 630, 235, 365),],
433       aes(x = year, y = latent_personalism, colour = gwf_country)) +
434     geom_step() + facet_grid(rows = vars(cowcode)) + blue_light +
435     labs(x = "Year", y = "Latent personalism") +
436     scale_color_manual(name='Country', labels = vars(gwf_country),
437                       values = palette("Okabe-Ito"))
438
439 #     Heatmap of personalism distribution per year
440 ggplot(data = autocracy_data, mapping = aes(y = latent_personalism,
441                                           x = year)) +
442     geom_bin2d(binwidth = c(1, 0.1)) + blue_light +
443     labs(y = "Latent personalism", x = "Year")
444
445 #     Calculation of country means
446 latent_personalism_full_countrymeans <- tibble(
447   countrycode = unique(autocracy_data$cowcode)[order(unique(
448     autocracy_data$cowcode))],
449   countrymean = tapply(autocracy_data$latent_personalism,
450                       autocracy_data$cowcode,
451                       mean)
452 )
453 #     Histogram and QQ-plot of country means
454 ggplot(data = latent_personalism_full_countrymeans, aes(x = countrymean)) +
455     geom_histogram(binwidth = 0.1, fill = "lightblue") +
456     labs(x = "Latent personalism", y = "Count") + blue_light
457 ggplot(data = latent_personalism_full_countrymeans, aes(sample = countrymean)) +
458     geom_qq(colour = "lightblue") + geom_qq_line() +
459     labs(x = "Reference normal distribution", y = "Latent personalism") +
460     blue_light
461
462 #     Compute summary statistics per year
463 latent_personalism_full_summary <-
464   tibble(year = sort(unique(autocracy_data$year)),
465         Mean = tapply(autocracy_data$latent_personalism,
466                     autocracy_data$year, mean),
467         Q1 = tapply(autocracy_data$latent_personalism,
468                   autocracy_data$year,
469                   quantile, prob = 0.25),
470         Median = tapply(autocracy_data$latent_personalism,
471                       autocracy_data$year, median),
472         Q3 = tapply(autocracy_data$latent_personalism,

```

```

473         autocracy_data$year,
474         quantile, prob = 0.75),
475     SD = tapply(autocracy_data$latent_personalism,
476               autocracy_data$year, sd),
477     Skew = tapply(autocracy_data$latent_personalism,
478                 autocracy_data$year, FUN = Skew),
479     Kurtosis = tapply(autocracy_data$latent_personalism,
480                      autocracy_data$year, FUN = Kurt) )
481
482 #     Pivot summary statistics table for use in plotting
483 latent_personalism_full_summary_longtable <-
484   pivot_longer(latent_personalism_full_summary, 2:8, names_to = "Statistic",
485               values_to = "Value")
486
487 #     Plot Mean, median, first quartile, third quartile
488 #     and standard deviation over time
489 withr::with_options(
490   list(ggplot2.discrete.colour = lines_palette),
491   print(ggplot(data = latent_personalism_full_summary_longtable[
492     latent_personalism_full_summary_longtable$Statistic %in% c("Mean", "Q1",
493     "Median", "Q3",
494     "SD"), ],
495     aes(x = year, y = Value, colour = Statistic)) +
496     geom_line() + blue_light + labs(x = "Year", y = "Latent personalism"))
497 )
498
499 #     Plot skew and kurtosis over time
500 ggplot(data = latent_personalism_full_summary_longtable[
501   latent_personalism_full_summary_longtable$Statistic %in%
502   c("Skew", "Kurtosis"), ],
503   aes(x = year, y = Value, linetype = Statistic)) +
504   geom_line() +
505   labs(x = "Year", y = "Latent personalism") +
506   blue_light
507
508 #     Complete cases
509
510 #     Heatmap of personalism distribution per year
511 ggplot(data = datacomplete, mapping = aes(y = latent_personalism,
512     x = year)) +
513   geom_bin2d(binwidth = c(1, 0.1)) + blue_light +
514   labs(y = "Latent personalism", x = "Year")
515
516 #     Calculation of country means
517 latent_personalism_complete_countrymeans <- tibble(
518   countrycode = unique(datacomplete$cowcode)[order(unique(
519     datacomplete$cowcode))],
520   countrymean = tapply(datacomplete$latent_personalism, datacomplete$cowcode,
521     mean)
522 )
523
524 #     Histogram and QQ-plot of country means
525 ggplot(data = latent_personalism_complete_countrymeans, aes(x = countrymean)) +
526   geom_histogram(binwidth = 0.1, fill = "lightblue") +
527   labs(x = "Latent personalism", y = "Count") + blue_light
528 ggplot(data = latent_personalism_complete_countrymeans,
529   aes(sample = countrymean)) +
530   geom_qq(colour = "lightblue") + geom_qq_line() +
531   labs(x = "Reference normal distribution", y = "Latent personalism") +

```

```

532   blue_light
533
534   #       Compute summary statistics per year
535   latent_personalism_complete_summary <-
536     tibble(year = unique(datacomplete$year)[order(unique(
537       datacomplete$year))],
538       Mean = tapply(datacomplete$latent_personalism, datacomplete$year,
539         mean),
540       Q1 = tapply(datacomplete$latent_personalism, datacomplete$year,
541         quantile, prob = 0.25),
542       Median = tapply(datacomplete$latent_personalism,
543         datacomplete$year, median),
544       Q3 = tapply(datacomplete$latent_personalism,
545         datacomplete$year,
546         quantile, prob = 0.75),
547       SD = tapply(datacomplete$latent_personalism,
548         datacomplete$year, sd),
549       Skew = tapply(datacomplete$latent_personalism,
550         datacomplete$year, FUN = Skew),
551       Kurtosis = tapply(datacomplete$latent_personalism,
552         datacomplete$year, FUN = Kurt) )
553
554   #       Pivot summary statistics table for use in plotting
555   latent_personalism_complete_summary_longtable <-
556     pivot_longer(latent_personalism_complete_summary, 2:8, names_to = "Statistic",
557       values_to = "Value")
558
559   #       Plot Mean, median, first quartile, third quartile
560   #       and standard deviation over time
561   withr::with_options(
562     list(ggplot2.discrete.colour = lines_palette),
563     print(ggplot(data = latent_personalism_complete_summary_longtable[
564       latent_personalism_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
565         "Median", "Q3",
566         "SD"), ],
567       aes(x = year, y = Value, colour = Statistic)) +
568       geom_line() + blue_light + labs(x = "Year", y = "Latent personalism"))
569   )
570
571   #       Plot skew and kurtosis over time
572   ggplot(data = latent_personalism_complete_summary_longtable[
573     latent_personalism_complete_summary_longtable$Statistic %in%
574     c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
575     geom_line() + labs(x = "Year", y = "Latent personalism") + blue_light
576
577   ##### Freedom of expression descriptives #####
578
579   #       Item description and scale evaluation
580
581   #       Government censorship effort
582
583   #       Establish observed range
584   range(autocracy_data$v2mecenefm)
585
586   #       Heatmap of government censorship effort distribution per year
587   ggplot(data = autocracy_data, mapping = aes(y = v2mecenefm, x = year)) +
588     geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
589     labs(y = "Government censorship effort", x = "Year")
590

```

```

591 # Calculation of country means
592 government_censorship_countrymeans <- tibble(
593   countrycode = unique(autocracy_data$cowcode)[order(unique(
594     autocracy_data$cowcode))],
595   countrymean = tapply(autocracy_data$v2mecenefm, autocracy_data$cowcode, mean)
596 )
597 # Histogram and QQ-plot of country means
598 ggplot(data = government_censorship_countrymeans, aes(x = countrymean)) +
599   geom_histogram(binwidth = 0.5, fill = "lightblue") +
600   labs(x = "Government censorship effort", y = "Count") + blue_light
601 ggplot(data = government_censorship_countrymeans, aes(sample = countrymean)) +
602   geom_qq(colour = "lightblue") + geom_qq_line() +
603   labs(x = "Reference normal distribution",
604     y = "Government censorship effort") + blue_light
605
606 # Compute summary statistics per year
607 government_censorship_summary <-
608   tibble(year = sort(unique(autocracy_data$year)),
609     Mean = tapply(autocracy_data$v2mecenefm, autocracy_data$year, mean),
610     Q1 = tapply(autocracy_data$v2mecenefm, autocracy_data$year,
611       quantile, prob = 0.25),
612     Median = tapply(autocracy_data$v2mecenefm, autocracy_data$year,
613       median),
614     Q3 = tapply(autocracy_data$v2mecenefm, autocracy_data$year,
615       quantile, prob = 0.75),
616     SD = tapply(autocracy_data$v2mecenefm, autocracy_data$year, sd),
617     Skew = tapply(autocracy_data$v2mecenefm, autocracy_data$year,
618       FUN = Skew),
619     Kurtosis = tapply(autocracy_data$v2mecenefm, autocracy_data$year,
620       FUN = Kurt) )
621
622 # Pivot summary statistics for use in plotting
623 government_censorship_summary_longtable <-
624   pivot_longer(government_censorship_summary, 2:8, names_to = "Statistic",
625     values_to = "Value")
626
627 # Plot Mean, median, first quartile, third quartile
628 # and standard deviation over time
629 withr::with_options(
630   list(ggplot2.discrete.colour = lines_palette),
631   print(ggplot(data = government_censorship_summary_longtable[
632     government_censorship_summary_longtable$Statistic %in% c("Mean", "Q1",
633       "Median", "Q3",
634       "SD"), ],
635     aes(x = year, y = Value, colour = Statistic)) + geom_line() +
636     blue_light + labs(x = "Year", y = "Government censorship effort"))
637 )
638
639 # Plot skew and kurtosis over time
640 ggplot(data = government_censorship_summary_longtable[
641   government_censorship_summary_longtable$Statistic %in%
642   c("Skew", "Kurtosis"), ],
643   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
644   labs(x = "Year", y = "Government censorship effort") + blue_light
645
646 # Harassment of journalists
647
648 # Establish observed range
649 range(autocracy_data$v2meharjrn)

```

```

650
651 # Heatmap of personalism distribution per year
652 ggplot(data = autocracy_data, mapping = aes(y = v2meharjrn, x = year)) +
653   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
654   labs(y = "Journalist harassment", x = "Year")
655
656 # Calculation of country means
657 journalist_harassment_countrymeans <- tibble(
658   countrycode = unique(autocracy_data$cowcode)[order(unique(
659     autocracy_data$cowcode))],
660   countrymean = tapply(autocracy_data$v2meharjrn, autocracy_data$cowcode,
661     mean)
662 )
663 # Histogram and QQ-plot of country means
664 ggplot(data = journalist_harassment_countrymeans, aes(x = countrymean)) +
665   geom_histogram(binwidth = 0.5, fill = "lightblue") +
666   labs(x = "Journalist harassment", y = "Count") + blue_light
667 ggplot(data = journalist_harassment_countrymeans, aes(sample = countrymean)) +
668   geom_qq(colour = "lightblue") +
669   geom_qq_line() +
670   labs(x = "Reference normal distribution", y = "Journalist harassment") +
671   blue_light
672
673 # Compute summary statistics per year
674 journalist_harassment_summary <-
675   tibble(year = sort(unique(autocracy_data$year)),
676     Mean = tapply(autocracy_data$v2meharjrn, autocracy_data$year, mean),
677     Q1 = tapply(autocracy_data$v2meharjrn, autocracy_data$year,
678       quantile, prob = 0.25),
679     Median = tapply(autocracy_data$v2meharjrn, autocracy_data$year, median),
680     Q3 = tapply(autocracy_data$v2meharjrn, autocracy_data$year,
681       quantile, prob = 0.75),
682     SD = tapply(autocracy_data$v2meharjrn, autocracy_data$year, sd),
683     Skew = tapply(autocracy_data$v2meharjrn, autocracy_data$year,
684       FUN = Skew),
685     Kurtosis = tapply(autocracy_data$v2meharjrn, autocracy_data$year,
686       FUN = Kurt) )
687
688 # Pivot summary statistics table for use in plotting
689 journalist_harassment_summary_longtable <-
690   pivot_longer(journalist_harassment_summary, 2:8, names_to = "Statistic",
691     values_to = "Value")
692
693 # Plot Mean, median, first quartile, third quartile
694 # and standard deviation over time
695 withr::with_options(
696   list(ggplot2.discrete.colour = lines_palette),
697   print(ggplot(data = journalist_harassment_summary_longtable[
698     journalist_harassment_summary_longtable$Statistic %in% c("Mean", "Q1",
699       "Median", "Q3",
700       "SD"), ],
701     aes(x = year, y = Value, colour = Statistic)) +
702     geom_line() + blue_light + labs(x = "Year",
703       y = "Journalist harassment"))
704 )
705
706 # Plot skew and kurtosis over time
707 ggplot(data = journalist_harassment_summary_longtable[
708   journalist_harassment_summary_longtable$Statistic %in%

```

```

709     c("Skew", "Kurtosis"), ],
710     aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
711     labs(x = "Year", y = "Journalist harassment") + blue_light
712
713 #     Media self-censorship
714
715 #     Establish observed range
716 range(autocracy_data$v2meslfcen)
717
718 #     Heatmap of media self-censorship distribution per year
719 ggplot(data = autocracy_data, mapping = aes(y = v2meslfcen,
720                                             x = year)) +
721     geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
722     labs(y = "Media self-censorship", x = "Year")
723
724 #     Calculation of country means
725 media_selfcensor_countrymeans <- tibble(
726     countrycode = unique(autocracy_data$cowcode)[order(unique(
727     autocracy_data$cowcode))],
728     countrymean = tapply(autocracy_data$v2meslfcen, autocracy_data$cowcode,
729     mean)
730 )
731 #     Histogram and QQ-plot of country means
732 ggplot(data = media_selfcensor_countrymeans, aes(x = countrymean)) +
733     geom_histogram(binwidth = 0.5, fill = "lightblue") +
734     labs(x = "Media self-censorship", y = "Count") + blue_light
735 ggplot(data = media_selfcensor_countrymeans, aes(sample = countrymean)) +
736     geom_qq(colour = "lightblue") + geom_qq_line() +
737     labs(x = "Reference normal distribution",
738     y = "Media self-censorship") + blue_light
739
740 #     Compute summary statistics per year
741 media_selfcensor_summary <-
742     tibble(year = sort(unique(autocracy_data$year)),
743     Mean = tapply(autocracy_data$v2meslfcen, autocracy_data$year, mean),
744     Q1 = tapply(autocracy_data$v2meslfcen, autocracy_data$year,
745     quantile, prob = 0.25),
746     Median = tapply(autocracy_data$v2meslfcen, autocracy_data$year,
747     median),
748     Q3 = tapply(autocracy_data$v2meslfcen, autocracy_data$year,
749     quantile, prob = 0.75),
750     SD = tapply(autocracy_data$v2meslfcen, autocracy_data$year, sd),
751     Skew = tapply(autocracy_data$v2meslfcen, autocracy_data$year,
752     FUN = Skew),
753     Kurtosis = tapply(autocracy_data$v2meslfcen, autocracy_data$year,
754     FUN = Kurt) )
755
756 #     Pivot summary statistics table for use in plotting
757 media_selfcensor_summary_longtable <-
758     pivot_longer(media_selfcensor_summary, 2:8, names_to = "Statistic",
759     values_to = "Value")
760
761 #     Plot Mean, median, first quartile, third quartile
762 #     and standard deviation over time
763 withr::with_options(
764     list(ggplot2.discrete.colour = lines_palette),
765     print(ggplot(data = media_selfcensor_summary_longtable[
766     media_selfcensor_summary_longtable$Statistic %in% c("Mean", "Q1",
767     "Median", "Q3",

```

```

768                                     "SD"), ],
769     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
770     labs(x = "Year", y = "Media self-censorship"))
771 )
772
773 #     Plot skew and kurtosis over time
774 ggplot(data = media_selfcensor_summary_longtable[
775     media_selfcensor_summary_longtable$Statistic %in%
776     c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
777     geom_line() + labs(x = "Year", y = "Media self-censorship") + blue_light
778
779 #     Freedom of discussion
780
781 #     Freedom of discussion for men
782
783 #     Establish observed range
784 range(autocracy_data$v2cldiscm)
785
786 #     Heatmap of media freedom of discussion for men distribution per year
787 ggplot(data = autocracy_data, mapping = aes(y = v2cldiscm, x = year)) +
788     geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
789     labs(y = "Freedom of discussion for men", x = "Year")
790
791 #     Calculation of country means
792 free_disc_men_countrymeans <- tibble(
793     countrycode = unique(autocracy_data$cowcode)[order(unique(
794     autocracy_data$cowcode))],
795     countrymean = tapply(autocracy_data$v2cldiscm, autocracy_data$cowcode,
796     mean)
797 )
798 #     Histogram and QQ-plot of country means
799 ggplot(data = free_disc_men_countrymeans, aes(x = countrymean)) +
800     geom_histogram(binwidth = 0.5, fill = "lightblue") +
801     labs(x = "Freedom of discussion for men", y = "Count") + blue_light
802 ggplot(data = free_disc_men_countrymeans, aes(sample = countrymean)) +
803     geom_qq(colour = "lightblue") + geom_qq_line() +
804     labs(x = "Reference normal distribution",
805     y = "Freedom of discussion for men") + blue_light
806
807 #     Compute summary statistics per year
808 free_disc_men_summary <-
809     tibble(year = sort(unique(autocracy_data$year)),
810     Mean = tapply(autocracy_data$v2cldiscm, autocracy_data$year, mean),
811     Q1 = tapply(autocracy_data$v2cldiscm, autocracy_data$year,
812     quantile, prob = 0.25),
813     Median = tapply(autocracy_data$v2cldiscm, autocracy_data$year, median),
814     Q3 = tapply(autocracy_data$v2cldiscm, autocracy_data$year,
815     quantile, prob = 0.75),
816     SD = tapply(autocracy_data$v2cldiscm, autocracy_data$year, sd),
817     Skew = tapply(autocracy_data$v2cldiscm, autocracy_data$year,
818     FUN = Skew),
819     Kurtosis = tapply(autocracy_data$v2cldiscm, autocracy_data$year,
820     FUN = Kurt) )
821
822 #     Pivot summary statistics table for use in plotting
823 free_disc_men_summary_longtable <-
824     pivot_longer(free_disc_men_summary, 2:8, names_to = "Statistic",
825     values_to = "Value")
826

```



```

827 #       Plot Mean, median, first quartile, third quartile
828 #       and standard deviation over time
829 withr::with_options(
830   list(ggplot2.discrete.colour = lines_palette),
831   print(ggplot(data = free_disc_men_summary_longtable[
832     free_disc_men_summary_longtable$Statistic %in% c("Mean", "Q1",
833     "Median", "Q3",
834     "SD"), ],
835     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
836     labs(x = "Year", y = "Freedom of discussion for men"))
837 )
838
839 #       Plot skew and kurtosis over time
840 ggplot(data = free_disc_men_summary_longtable[
841   free_disc_men_summary_longtable$Statistic %in%
842   c("Skew", "Kurtosis"), ],
843   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
844   labs(x = "Year", y = "Freedom of discussion for men") + blue_light
845
846 #       Freedom of discussion for women
847
848 #       Establish observed range
849 range(autocracy_data$v2cldiscw)
850
851 #       Heatmap of freedom of discussion for women distribution per year
852 ggplot(data = autocracy_data, mapping = aes(y = v2cldiscw, x = year)) +
853   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
854   labs(y = "Freedom of discussion for women", x = "Year")
855
856 #       Calculation of country means
857 free_disc_women_countrymeans <- tibble(
858   countrycode = unique(autocracy_data$cowcode)[order(unique(
859     autocracy_data$cowcode))],
860   countrymean = tapply(autocracy_data$v2cldiscw,
861     autocracy_data$cowcode,
862     mean)
863 )
864 #       Histogram and QQ-plot of country means
865 ggplot(data = free_disc_women_countrymeans, aes(x = countrymean)) +
866   geom_histogram(binwidth = 0.5, fill = "lightblue") +
867   labs(x = "Freedom of discussion for women", y = "Count") + blue_light
868 ggplot(data = free_disc_women_countrymeans, aes(sample = countrymean)) +
869   geom_qq(colour = "lightblue") + geom_qq_line() +
870   labs(x = "Reference normal distribution",
871     y = "Freedom of discussion for women") + blue_light
872
873 #       Compute summary statistics per year
874 free_disc_women_summary <-
875   tibble(year = sort(unique(autocracy_data$year)),
876     Mean = tapply(autocracy_data$v2cldiscw, autocracy_data$year, mean),
877     Q1 = tapply(autocracy_data$v2cldiscw, autocracy_data$year,
878     quantile, prob = 0.25),
879     Median = tapply(autocracy_data$v2cldiscw, autocracy_data$year, median),
880     Q3 = tapply(autocracy_data$v2cldiscw, autocracy_data$year,
881     quantile, prob = 0.75),
882     SD = tapply(autocracy_data$v2cldiscw, autocracy_data$year, sd),
883     Skew = tapply(autocracy_data$v2cldiscw, autocracy_data$year,
884     FUN = Skew),
885     Kurtosis = tapply(autocracy_data$v2cldiscw, autocracy_data$year,

```

```

886             FUN = Kurt) )
887
888 #       Pivot summary statistics table for use in plotting
889 free_disc_women_summary_longtable <-
890   pivot_longer(free_disc_women_summary, 2:8, names_to = "Statistic",
891               values_to = "Value")
892
893 #       Plot Mean, median, first quartile, third quartile
894 #       and standard deviation over time
895 withr::with_options(
896   list(ggplot2.discrete.colour = lines_palette),
897   print(ggplot(data = free_disc_women_summary_longtable[
898     free_disc_women_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
899     "Q3", "SD"), ],
900     aes(x = year, y = Value, colour = Statistic)) +
901     geom_line() + blue_light + labs(x = "Year",
902     y = "Freedom of discussion for women"))
903 )
904
905 #       Plot skew and kurtosis over time
906 ggplot(data = free_disc_women_summary_longtable[
907   free_disc_women_summary_longtable$Statistic %in%
908   c("Skew", "Kurtosis"), ],
909   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
910   labs(x = "Year", y = "Freedom of discussion for women") + blue_light
911
912 #       Sub-item intercorrelation and freedom of discussion description
913
914 #       Calculate intercorrelation by year
915 free_disc_cors <-
916   by(data = autocracy_data[, c("v2cldiscm", "v2cldiscw")],
917     INDICES = autocracy_data$year, FUN = cor)
918
919 #       Create table for correlations
920 table_free_disc_cors <- tibble(Year = sort(unique(autocracy_data$year)),
921   "Correlation" = NA)
922
923 #       Fill table
924 for (y in sort(unique(autocracy_data$year))) {
925   table_free_disc_cors[table_free_disc_cors$Year == y,
926     "Correlation"] <-
927     free_disc_cors[[as.character(y)]] [2,1]
928 }
929
930 #       Plot correlations by year
931 ggplot(table_free_disc_cors, aes(x = Year, y = Correlation)) +
932   geom_line(colour = "tomato") + blue_light
933
934 #       Establish range of correlations
935 range(table_free_disc_cors$"Correlation")
936
937 #       Establish observed range for scale
938 range(autocracy_data$v2xcl_disc)
939
940 #       Heatmap of freedom of discussion distribution per year
941 ggplot(data = autocracy_data, mapping = aes(y = v2xcl_disc, x = year)) +
942   geom_bin2d(binwidth = c(1, 0.1)) + blue_light +
943   labs(y = "Freedom of discussion", x = "Year")
944

```

```

945 # Calculation of country means
946 free_disc_countrymeans <- tibble(
947   countrycode = unique(autocracy_data$cowcode)[order(unique(
948     autocracy_data$cowcode))],
949   countrymean = tapply(autocracy_data$v2xcl_disc, autocracy_data$cowcode,
950     mean)
951 )
952 # Histogram and QQ-plot of country means
953 ggplot(data = free_disc_countrymeans, aes(x = countrymean)) +
954   geom_histogram(binwidth = 0.1, fill = "lightblue") +
955   labs(x = "Freedom of discussion", y = "Count") + blue_light
956 ggplot(data = free_disc_countrymeans, aes(sample = countrymean)) +
957   geom_qq(colour = "lightblue") + geom_qq_line() +
958   labs(x = "Reference normal distribution",
959     y = "Freedom of discussion") + blue_light
960
961 # Compute scale summary statistics per year
962 free_disc_summary <-
963   tibble(year = sort(unique(autocracy_data$year)),
964     Mean = tapply(autocracy_data$v2xcl_disc, autocracy_data$year, mean),
965     Q1 = tapply(autocracy_data$v2xcl_disc, autocracy_data$year,
966       quantile, prob = 0.25),
967     Median = tapply(autocracy_data$v2xcl_disc, autocracy_data$year,
968       median),
969     Q3 = tapply(autocracy_data$v2xcl_disc, autocracy_data$year,
970       quantile, prob = 0.75),
971     SD = tapply(autocracy_data$v2xcl_disc, autocracy_data$year, sd),
972     Skew = tapply(autocracy_data$v2xcl_disc, autocracy_data$year,
973       FUN = Skew),
974     Kurtosis = tapply(autocracy_data$v2xcl_disc, autocracy_data$year,
975       FUN = Kurt) )
976
977 # Pivot summary statistics table for use in plotting
978 free_disc_summary_longtable <-
979   pivot_longer(free_disc_summary, 2:8, names_to = "Statistic",
980     values_to = "Value")
981
982 # Plot mean, median, first quartile, third quartile
983 # and standard deviation of scale over time
984 withr::with_options(
985   list(ggplot2.discrete.colour = lines_palette),
986   print(ggplot(data = free_disc_summary_longtable[
987     free_disc_summary_longtable$Statistic %in% c("Mean", "Q1", "Median", "Q3",
988       "SD"), ],
989     aes(x = year, y = Value, colour = Statistic)) +
990     geom_line() + blue_light + labs(x = "Year",
991       y = "Freedom of discussion"))
992 )
993
994 # Plot skew and kurtosis of scale over time
995 ggplot(data = free_disc_summary_longtable[
996   free_disc_summary_longtable$Statistic %in%
997   c("Skew", "Kurtosis"), ],
998   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
999   labs(x = "Year", y = "Freedom of discussion") + blue_light
1000
1001
1002 # Freedom of academic and cultural expression
1003

```

```

1004 # Establish observed range
1005 range(autocracy_data$v2clacfree)
1006
1007 # Heatmap of media self-censorship distribution per year
1008 ggplot(data = autocracy_data, mapping = aes(y = v2clacfree, x = year)) +
1009   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1010   labs(y = "Freedom of acad. and cult. expr.", x = "Year")
1011
1012 # Calculation of country means
1013 free_acad_cult_expr_countrymeans <- tibble(
1014   countrycode = unique(autocracy_data$cowcode)[order(unique(
1015     autocracy_data$cowcode))],
1016   countrymean = tapply(autocracy_data$v2clacfree, autocracy_data$cowcode,
1017     mean)
1018 )
1019 # Histogram and QQ-plot of country means
1020 ggplot(data = free_acad_cult_expr_countrymeans, aes(x = countrymean)) +
1021   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1022   labs(x = "Freedom of acad. and cult. expr.", y = "Count") + blue_light
1023 ggplot(data = free_acad_cult_expr_countrymeans, aes(sample = countrymean)) +
1024   geom_qq(colour = "lightblue") + geom_qq_line() +
1025   labs(x = "Reference normal distribution",
1026     y = "Freedom of acad. and cult. expr.") + blue_light
1027
1028 # Compute summary statistics per year
1029 free_acad_cult_expr_summary <-
1030   tibble(year = sort(unique(autocracy_data$year)),
1031     Mean = tapply(autocracy_data$v2clacfree, autocracy_data$year, mean),
1032     Q1 = tapply(autocracy_data$v2clacfree, autocracy_data$year,
1033       quantile, prob = 0.25),
1034     Median = tapply(autocracy_data$v2clacfree, autocracy_data$year,
1035       median),
1036     Q3 = tapply(autocracy_data$v2clacfree, autocracy_data$year,
1037       quantile, prob = 0.75),
1038     SD = tapply(autocracy_data$v2clacfree, autocracy_data$year, sd),
1039     Skew = tapply(autocracy_data$v2clacfree, autocracy_data$year,
1040       FUN = Skew),
1041     Kurtosis = tapply(autocracy_data$v2clacfree, autocracy_data$year,
1042       FUN = Kurt) )
1043
1044 # Pivot summary statistics table for use in plotting
1045 free_acad_cult_expr_summary_longtable <-
1046   pivot_longer(free_acad_cult_expr_summary, 2:8, names_to = "Statistic",
1047     values_to = "Value")
1048
1049 # Plot Mean, median, first quartile, third quartile
1050 # and standard deviation over time
1051 withr::with_options(
1052   list(ggplot2.discrete.colour = lines_palette),
1053   print(ggplot(data = free_acad_cult_expr_summary_longtable[
1054     free_acad_cult_expr_summary_longtable$Statistic %in% c("Mean", "Q1",
1055       "Median", "Q3",
1056       "SD"), ],
1057     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
1058     labs(x = "Year", y = "Freedom of acad. and cult. expr.))
1059 )
1060
1061 # Plot skew and kurtosis over time
1062 ggplot(data = free_acad_cult_expr_summary_longtable[

```

```

1063   free_acad_cult_expr_summary_longtable$Statistic %in%
1064     c("Skew", "Kurtosis"), ],
1065   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1066   labs(x = "Year", y = "Freedom of acad. and cult. expr.") + blue_light
1067
1068 #   Reliability analysis
1069
1070 #   Calculate Cronbach's Alpha by year
1071 free_expr_cronbachalphas <-
1072   by(data = autocracy_data[, c("v2mecenefm", "v2meharjrn", "v2meslfcen",
1073     "v2xcl_disc", "v2clacfree")],
1074     INDICES = autocracy_data$year, FUN = CronbachAlpha, cond = TRUE)
1075
1076 #   Create table for Cronbach's Alpha values
1077 table_free_expr_cronbachalphas <-
1078   tibble(year = sort(unique(autocracy_data$year)),
1079     "All items" = NA, "Government censorship effort" = NA,
1080     "Journalist harassment" = NA, "Media self-censorship" = NA,
1081     "Freedom of discussion" = NA, "Freedom of acad. and cult. expr." = NA
1082   )
1083
1084 #   Fill in table
1085 for (y in sort(unique(autocracy_data$year))) {
1086   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1087     "All items"] <-
1088     free_expr_cronbachalphas[[as.character(y)]]["unconditional"]
1089   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1090     "Government censorship effort"] <-
1091     free_expr_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1092     ][1, "Cronbach Alpha"]
1093   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1094     "Journalist harassment"] <-
1095     free_expr_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1096     ][2, "Cronbach Alpha"]
1097   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1098     "Media self-censorship"] <-
1099     free_expr_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1100     ][3, "Cronbach Alpha"]
1101   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1102     "Freedom of discussion"] <-
1103     free_expr_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1104     ][4, "Cronbach Alpha"]
1105   table_free_expr_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1106     "Freedom of acad. and cult. expr."] <-
1107     free_expr_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1108     ][5, "Cronbach Alpha"]
1109 }
1110
1111 #   Pivot table for use in plotting
1112 longtab_free_expr_cronbachalphas <-
1113   pivot_longer(table_free_expr_cronbachalphas, 2:7, names_to = "Type",
1114     values_to = "Value")
1115
1116 #   Plot Cronbach's Alpha by year
1117 ggplot(longtab_free_expr_cronbachalphas, aes(x = year, y = Value,
1118     colour = Type)) +
1119   geom_line() + blue_light
1120
1121 #   Calculate range of yearly Cronbach's Alpha

```

```

1122 range(table_free_expr_cronbachalphas$"All items")
1123
1124 # Original scale
1125
1126 # Heatmap of freedom of expression distribution per year
1127 ggplot(data = autocracy_data, mapping = aes(y = v2x_freexp, x = year)) +
1128   geom_bin2d(binwidth = c(1, 0.1)) + blue_light +
1129   labs(y = "Freedom of expression", x = "Year")
1130
1131 # Calculation of country means
1132 free_expr_unscaled_countrymeans <- tibble(
1133   countrycode = unique(autocracy_data$cowcode)[order(unique(
1134     autocracy_data$cowcode))],
1135   countrymean = tapply(autocracy_data$v2x_freexp, autocracy_data$cowcode,
1136     mean)
1137 )
1138 # Histogram and QQ-plot of country means
1139 ggplot(data = free_expr_unscaled_countrymeans, aes(x = countrymean)) +
1140   geom_histogram(binwidth = 0.1, fill = "lightblue") +
1141   labs(x = "Freedom of expression", y = "Count") + blue_light
1142 ggplot(data = free_expr_unscaled_countrymeans, aes(sample = countrymean)) +
1143   geom_qq(colour = "lightblue") + geom_qq_line() +
1144   labs(x = "Reference normal distribution", y = "Freedom of expression") +
1145   blue_light
1146
1147 # Compute summary statistics per year
1148 free_expr_unscaled_summary <-
1149   tibble(year = sort(unique(autocracy_data$year)),
1150     Mean = tapply(autocracy_data$v2x_freexp, autocracy_data$year, mean),
1151     Q1 = tapply(autocracy_data$v2x_freexp, autocracy_data$year,
1152       quantile, prob = 0.25),
1153     Median = tapply(autocracy_data$v2x_freexp, autocracy_data$year,
1154       median),
1155     Q3 = tapply(autocracy_data$v2x_freexp, autocracy_data$year,
1156       quantile, prob = 0.75),
1157     SD = tapply(autocracy_data$v2x_freexp, autocracy_data$year, sd),
1158     Skew = tapply(autocracy_data$v2x_freexp, autocracy_data$year,
1159       FUN = Skew),
1160     Kurtosis = tapply(autocracy_data$v2x_freexp, autocracy_data$year,
1161       FUN = Kurt) )
1162
1163 # Pivot summary statistics table for use in plotting
1164 free_expr_unscaled_summary_longtable <-
1165   pivot_longer(free_expr_unscaled_summary, 2:8, names_to = "Statistic",
1166     values_to = "Value")
1167
1168 # Plot Mean, median, first quartile, third quartile
1169 # and standard deviation over time
1170 withr::with_options(
1171   list(ggplot2.discrete.colour = lines_palette),
1172   print(ggplot(data = free_expr_unscaled_summary_longtable[
1173     free_expr_unscaled_summary_longtable$Statistic %in% c("Mean", "Q1",
1174       "Median", "Q3",
1175       "SD"), ],
1176     aes(x = year, y = Value, colour = Statistic)) +
1177     geom_line() + blue_light + labs(x = "Year", y = "Freedom of expression"))
1178 )
1179
1180 # Plot skew and kurtosis over time

```

```

1181 ggplot(data = free_expr_unscaled_summary_longtable[
1182   free_expr_unscaled_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
1183   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1184   labs(x = "Year", y = "Freedom of expression") + blue_light
1185
1186
1187 #   Final scale descriptives
1188
1189 #   Entire dataset
1190
1191 #   Heatmap of freedom of expression distribution per year
1192 ggplot(data = autocracy_data, mapping = aes(y = free_expr_x100, x = year)) +
1193   geom_bin2d(binwidth = c(1, 10)) + blue_light +
1194   labs(y = "Freedom of expression", x = "Year")
1195
1196 #   Calculation of country means
1197 free_expr_full_countrymeans <- tibble(
1198   countrycode = unique(autocracy_data$cowcode)[order(unique(
1199     autocracy_data$cowcode))],
1200   countrymean = tapply(autocracy_data$free_expr_x100, autocracy_data$cowcode,
1201     mean)
1202 )
1203 #   Histogram and QQ-plot of country means
1204 ggplot(data = free_expr_full_countrymeans, aes(x = countrymean)) +
1205   geom_histogram(binwidth = 10, fill = "lightblue") +
1206   labs(x = "Freedom of expression", y = "Count") + blue_light
1207 ggplot(data = free_expr_full_countrymeans, aes(sample = countrymean)) +
1208   geom_qq(colour = "lightblue") + geom_qq_line() +
1209   labs(x = "Reference normal distribution", y = "Freedom of expression") +
1210   blue_light
1211
1212 #   Compute summary statistics per year
1213 free_expr_full_summary <-
1214   tibble(year = sort(unique(autocracy_data$year)),
1215     Mean = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1216       mean),
1217     Q1 = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1218       quantile, prob = 0.25),
1219     Median = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1220       median),
1221     Q3 = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1222       quantile, prob = 0.75),
1223     SD = tapply(autocracy_data$free_expr_x100, autocracy_data$year, sd),
1224     Skew = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1225       FUN = Skew),
1226     Kurtosis = tapply(autocracy_data$free_expr_x100, autocracy_data$year,
1227       FUN = Kurt) )
1228
1229 #   Pivot summary statistics table for use in plotting
1230 free_expr_full_summary_longtable <-
1231   pivot_longer(free_expr_full_summary, 2:8, names_to = "Statistic",
1232     values_to = "Value")
1233
1234 #   Plot Mean, median, first quartile, third quartile
1235 #   and standard deviation over time
1236 withr::with_options(
1237   list(ggplot2.discrete.colour = lines_palette),
1238   print(ggplot(data = free_expr_full_summary_longtable[
1239     free_expr_full_summary_longtable$Statistic %in% c("Mean", "Q1",

```

```

1240                                     "Median", "Q3",
1241                                     "SD"), ],
1242     aes(x = year, y = Value, colour = Statistic)) +
1243     geom_line() + blue_light + labs(x = "Year", y = "Freedom of expression"))
1244 )
1245
1246 #     Plot skew and kurtosis over time
1247 ggplot(data = free_expr_full_summary_longtable[
1248     free_expr_full_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
1249     aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1250     labs(x = "Year", y = "Freedom of expression") + blue_light
1251
1252 #     Complete cases
1253
1254 #     Heatmap of freedom of expression distribution per year
1255 ggplot(data = datacomplete, mapping = aes(y = free_expr_x100, x = year)) +
1256     geom_bin2d(binwidth = c(1, 10)) + blue_light +
1257     labs(y = "Freedom of expression", x = "Year")
1258
1259 #     Calculation of country means
1260 free_expr_complete_countrymeans <- tibble(
1261     countrycode = unique(datacomplete$cowcode)[order(unique(
1262     datacomplete$cowcode))],
1263     countrymean = tapply(datacomplete$free_expr_x100,
1264     datacomplete$cowcode, mean)
1265 )
1266
1267 #     Histogram and QQ-plot of country means
1268 ggplot(data = free_expr_complete_countrymeans, aes(x = countrymean)) +
1269     geom_histogram(binwidth = 10, fill = "lightblue") +
1270     labs(x = "Freedom of expression", y = "Count") + blue_light
1271 ggplot(data = free_expr_complete_countrymeans,
1272     aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
1273     geom_qq_line() + labs(x = "Reference normal distribution",
1274     y = "Freedom of expression") + blue_light
1275
1276 #     Compute summary statistics per year
1277 free_expr_complete_summary <-
1278     tibble(year = unique(datacomplete$year)[order(unique(
1279     datacomplete$year))],
1280     Mean = tapply(datacomplete$free_expr_x100, datacomplete$year, mean),
1281     Q1 = tapply(datacomplete$free_expr_x100, datacomplete$year,
1282     quantile, prob = 0.25),
1283     Median = tapply(datacomplete$free_expr_x100, datacomplete$year, median),
1284     Q3 = tapply(datacomplete$free_expr_x100, datacomplete$year,
1285     quantile, prob = 0.75),
1286     SD = tapply(datacomplete$free_expr_x100, datacomplete$year, sd),
1287     Skew = tapply(datacomplete$free_expr_x100, datacomplete$year, FUN = Skew),
1288     Kurtosis = tapply(datacomplete$free_expr_x100, datacomplete$year,
1289     FUN = Kurt) )
1290
1291 #     Pivot summary statistics table for use in plotting
1292 free_expr_complete_summary_longtable <-
1293     pivot_longer(free_expr_complete_summary, 2:8, names_to = "Statistic",
1294     values_to = "Value")
1295
1296 #     Plot Mean, median, first quartile, third quartile
1297 #     and standard deviation over time
1298 withr::with_options(

```



```

1299   list(ggplot2.discrete.colour = lines_palette),
1300   print(ggplot(data = free_expr_complete_summary_longtable[
1301     free_expr_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
1302     "Median", "Q3",
1303     "SD"), ],
1304     aes(x = year, y = Value, colour = Statistic)) +
1305     geom_line() + blue_light + labs(x = "Year", y = "Freedom of expression"))
1306 )
1307
1308 #   Plot skew and kurtosis over time
1309 ggplot(data = free_expr_complete_summary_longtable[
1310   free_expr_complete_summary_longtable$Statistic %in%
1311   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1312   geom_line() + labs(x = "Year", y = "Freedom of expression") + blue_light
1313
1314 ##### Freedom of assembly descriptives #####
1315
1316 # Entire dataset
1317
1318 # Establish observed range
1319 range(autocracy_data$v2caassemb, na.rm = TRUE)
1320
1321 # Heatmap of the original freedom of assembly variable
1322 # distribution per year
1323 ggplot(data = autocracy_data, mapping = aes(y = v2caassemb, x = year)) +
1324   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1325   labs(y = "Freedom of assembly", x = "Year")
1326
1327 # Calculation of country means
1328 free_assemb_org_countrymeans <- tibble(
1329   countrycode = unique(autocracy_data$cowcode)[order(unique(
1330     autocracy_data$cowcode))],
1331   countrymean = tapply(autocracy_data$v2caassemb, autocracy_data$cowcode,
1332     mean, na.rm = TRUE)
1333 )
1334
1335 # Histogram and QQ-plot of country means
1336 ggplot(data = free_assemb_org_countrymeans, aes(x = countrymean)) +
1337   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1338   labs(x = "Freedom of assembly", y = "Count") + blue_light
1339 ggplot(data = free_assemb_org_countrymeans,
1340   aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
1341   geom_qq_line() + labs(x = "Reference normal distribution",
1342     y = "Freedom of assembly") + blue_light
1343
1344 # Compute summary statistics per year
1345 free_assemb_org_summary <-
1346   tibble(year = unique(autocracy_data$year)[order(unique(
1347     autocracy_data$year))],
1348     Mean = tapply(autocracy_data$v2caassemb, autocracy_data$year, mean,
1349       na.rm = TRUE),
1350     Q1 = tapply(autocracy_data$v2caassemb, autocracy_data$year,
1351       quantile, prob = 0.25, na.rm = TRUE),
1352     Median = tapply(autocracy_data$v2caassemb, autocracy_data$year, median,
1353       na.rm = TRUE),
1354     Q3 = tapply(autocracy_data$v2caassemb, autocracy_data$year,
1355       quantile, prob = 0.75, na.rm = TRUE),
1356     SD = tapply(autocracy_data$v2caassemb, autocracy_data$year, sd,
1357       na.rm = TRUE),

```

```

1358     Skew = tapply(autocracy_data$v2caassemb, autocracy_data$year, FUN = Skew,
1359                   na.rm = TRUE),
1360     Kurtosis = tapply(autocracy_data$v2caassemb, autocracy_data$year,
1361                       FUN = Kurt, na.rm = TRUE) )
1362
1363 #       Pivot summary statistics table for use in plotting
1364 free_assemb_org_summary_longtable <-
1365   pivot_longer(free_assemb_org_summary, 2:8, names_to = "Statistic",
1366               values_to = "Value")
1367
1368 #       Plot Mean, median, first quartile, third quartile
1369 #       and standard deviation over time
1370 withr::with_options(
1371   list(ggplot2.discrete.colour = lines_palette),
1372   print(ggplot(data = free_assemb_org_summary_longtable[
1373     free_assemb_org_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
1374     "Q3", "SD"), ],
1375     aes(x = year, y = Value, colour = Statistic)) +
1376     geom_line() + blue_light + labs(x = "Year", y = "Freedom of assembly"))) )
1377
1378 #       Plot skew and kurtosis over time
1379 ggplot(data = free_assemb_org_summary_longtable[
1380   free_assemb_org_summary_longtable$Statistic %in%
1381   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1382   geom_line() + labs(x = "Year", y = "Freedom of assembly") + blue_light
1383
1384 # Complete cases
1385
1386 # Heatmap of freedom of assembly distribution per year
1387 ggplot(data = datacomplete, mapping = aes(y = v2caassemb, x = year)) +
1388   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1389   labs(y = "Freedom of assembly", x = "Year")
1390
1391 # Calculation of country means
1392 free_assemb_complete_countrymeans <- tibble(
1393   countrycode = unique(datacomplete$cowcode)[order(unique(
1394     datacomplete$cowcode))],
1395   countrymean = tapply(datacomplete$v2caassemb, datacomplete$cowcode, mean) )
1396
1397 # Histogram and QQ-plot of country means
1398 ggplot(data = free_assemb_complete_countrymeans, aes(x = countrymean)) +
1399   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1400   labs(x = "Freedom of assembly", y = "Count") + blue_light
1401 ggplot(data = free_assemb_complete_countrymeans,
1402   aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
1403   geom_qq_line() + labs(x = "Reference normal distribution",
1404   y = "Freedom of assembly") + blue_light
1405
1406 # Compute summary statistics per year
1407 free_assemb_complete_summary <-
1408   tibble(year = unique(datacomplete$year)[order(unique(
1409     datacomplete$year))],
1410   Mean = tapply(datacomplete$v2caassemb, datacomplete$year, mean),
1411   Q1 = tapply(datacomplete$v2caassemb, datacomplete$year,
1412     quantile, prob = 0.25),
1413   Median = tapply(datacomplete$v2caassemb, datacomplete$year, median),
1414   Q3 = tapply(datacomplete$v2caassemb, datacomplete$year,
1415     quantile, prob = 0.75),
1416   SD = tapply(datacomplete$v2caassemb, datacomplete$year, sd),

```

```

1417     Skew = tapply(datacomplete$v2caassemb, datacomplete$year, FUN = Skew),
1418     Kurtosis = tapply(datacomplete$v2caassemb, datacomplete$year, FUN = Kurt) )
1419
1420 #       Pivot summary statistics table for use in plotting
1421 free_assemb_complete_summary_longtable <-
1422   pivot_longer(free_assemb_complete_summary, 2:8, names_to = "Statistic",
1423               values_to = "Value")
1424
1425 # Plot Mean, median, first quartile, third quartile and standard deviation
1426 # over time
1427 withr::with_options(
1428   list(ggplot2.discrete.colour = lines_palette),
1429   print(ggplot(data = free_assemb_complete_summary_longtable[
1430     free_assemb_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
1431                                                           "Median", "Q3",
1432                                                           "SD"), ],
1433             aes(x = year, y = Value, colour = Statistic)) +
1434             geom_line() + blue_light + labs(x = "Year", y = "Freedom of assembly")) )
1435
1436 # Plot skew and kurtosis over time
1437 ggplot(data = free_assemb_complete_summary_longtable[
1438   free_assemb_complete_summary_longtable$Statistic %in%
1439   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1440   geom_line() + labs(x = "Year", y = "Freedom of assembly") + blue_light
1441
1442 ##### Freedom of movement descriptives #####
1443
1444 # Item descriptives
1445
1446 # Freedom of foreign movement
1447
1448 # Establish the observed range for the variable
1449 range(autocracy_data$v2clfmov)
1450
1451 # Heatmap of freedom of foreign movement per year
1452 ggplot(data = autocracy_data, mapping = aes(y = v2clfmov,
1453                                             x = year)) +
1454   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1455   labs(y = "Freedom of foreign movement", x = "Year")
1456
1457 # Calculation of country means
1458 free_move_foreign_countrymeans <- tibble(
1459   countrycode = unique(autocracy_data$cowcode)[order(unique(
1460     autocracy_data$cowcode))],
1461   countrymean = tapply(autocracy_data$v2clfmov, autocracy_data$cowcode,
1462                       mean) )
1463
1464 # Histogram and QQ-plot of country means
1465 ggplot(data = free_move_foreign_countrymeans, aes(x = countrymean)) +
1466   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1467   labs(x = "Freedom of foreign movement", y = "Count") + blue_light
1468 ggplot(data = free_move_foreign_countrymeans, aes(sample = countrymean)) +
1469   geom_qq(colour = "lightblue") + geom_qq_line() +
1470   labs(x = "Reference normal distribution", y = "Freedom of foreign movement") +
1471   blue_light
1472
1473 # Compute summary statistics per year
1474 free_move_foreign_summary <-
1475   tibble(year = sort(unique(autocracy_data$year)),

```

```

1476 Mean = tapply(autocracy_data$v2clfmov, autocracy_data$year, mean),
1477 Q1 = tapply(autocracy_data$v2clfmov, autocracy_data$year,
1478           quantile, prob = 0.25),
1479 Median = tapply(autocracy_data$v2clfmov, autocracy_data$year, median),
1480 Q3 = tapply(autocracy_data$v2clfmov, autocracy_data$year,
1481           quantile, prob = 0.75),
1482 SD = tapply(autocracy_data$v2clfmov, autocracy_data$year, sd),
1483 Skew = tapply(autocracy_data$v2clfmov, autocracy_data$year,
1484             FUN = Skew),
1485 Kurtosis = tapply(autocracy_data$v2clfmov, autocracy_data$year,
1486                 FUN = Kurt) )
1487
1488 # Pivot summary statistics table for use in plotting
1489 free_move_foreign_summary_longtable <-
1490   pivot_longer(free_move_foreign_summary, 2:8, names_to = "Statistic",
1491               values_to = "Value")
1492
1493 # Plot Mean, median, first quartile, third quartile
1494 # and standard deviation over time
1495 withr::with_options(
1496   list(ggplot2.discrete.colour = lines_palette),
1497   print(ggplot(data = free_move_foreign_summary_longtable[
1498     free_move_foreign_summary_longtable$Statistic %in% c("Mean", "Q1",
1499                                                         "Median", "Q3",
1500                                                         "SD"), ],
1501     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
1502     labs(x = "Year", y = "Freedom of foreign movement")) )
1503
1504 # Plot skew and kurtosis over time
1505 ggplot(data = free_move_foreign_summary_longtable[
1506   free_move_foreign_summary_longtable$Statistic %in%
1507   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1508   geom_line() + labs(x = "Year", y = "Freedom of foreign movement") + blue_light
1509
1510 # Freedom of domestic movement for men
1511
1512 # Establish the observed range for the variable
1513 range(autocracy_data$v2cldmovem)
1514
1515 # Heatmap of freedom of domestic movement for men distribution per year
1516 ggplot(data = autocracy_data, mapping = aes(y = v2cldmovem, x = year)) +
1517   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1518   labs(y = "Freedom of domestic movement for men", x = "Year")
1519
1520 # Calculation of country means
1521 free_move_men_countrymeans <- tibble(
1522   countrycode = unique(autocracy_data$cowcode)[order(unique(
1523     autocracy_data$cowcode))],
1524   countrymean = tapply(autocracy_data$v2cldmovem, autocracy_data$cowcode,
1525                       mean)
1526 )
1527 # Histogram and QQ-plot of country means
1528 ggplot(data = free_move_men_countrymeans, aes(x = countrymean)) +
1529   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1530   labs(x = "Freedom of domestic movement for men", y = "Count") + blue_light
1531 ggplot(data = free_move_men_countrymeans, aes(sample = countrymean)) +
1532   geom_qq(colour = "lightblue") + geom_qq_line() +
1533   labs(x = "Reference normal distribution",
1534        y = "Freedom of domestic movement for men") + blue_light

```

```

1535
1536 #       Compute summary statistics per year
1537 free_move_men_summary <-
1538   tibble(year = sort(unique(autocracy_data$year)),
1539         Mean = tapply(autocracy_data$v2cldmovem, autocracy_data$year, mean),
1540         Q1 = tapply(autocracy_data$v2cldmovem, autocracy_data$year,
1541                   quantile, prob = 0.25),
1542         Median = tapply(autocracy_data$v2cldmovem, autocracy_data$year,
1543                       median),
1544         Q3 = tapply(autocracy_data$v2cldmovem, autocracy_data$year,
1545                   quantile, prob = 0.75),
1546         SD = tapply(autocracy_data$v2cldmovem, autocracy_data$year, sd),
1547         Skew = tapply(autocracy_data$v2cldmovem, autocracy_data$year,
1548                     FUN = Skew),
1549         Kurtosis = tapply(autocracy_data$v2cldmovem, autocracy_data$year,
1550                          FUN = Kurt) )
1551
1552 #       Pivot summary statistics table for use in plotting
1553 free_move_men_summary_longtable <-
1554   pivot_longer(free_move_men_summary, 2:8, names_to = "Statistic",
1555               values_to = "Value")
1556
1557 #       Plot Mean, median, first quartile, third quartile
1558 #       and standard deviation over time
1559 withr::with_options(
1560   list(ggplot2.discrete.colour = lines_palette),
1561   print(ggplot(data = free_move_men_summary_longtable[
1562     free_move_men_summary_longtable$Statistic %in% c("Mean", "Q1",
1563                                                     "Median", "Q3",
1564                                                     "SD"), ],
1565     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
1566     labs(x = "Year", y = "Freedom of domestic movement for men")) )
1567
1568 #       Plot skew and kurtosis over time
1569 ggplot(data = free_move_men_summary_longtable[
1570   free_move_men_summary_longtable$Statistic %in%
1571   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1572   geom_line() + labs(x = "Year", y = "Freedom of domestic movement for men") +
1573   blue_light
1574
1575 #       Freedom of domestic movement for women
1576
1577 #       Establish the observed range for the variable
1578 range(autocracy_data$v2cldmovem)
1579
1580 #       Heatmap of freedom of domestic movement for women distribution per year
1581 ggplot(data = autocracy_data, mapping = aes(y = v2cldmovem, x = year)) +
1582   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1583   labs(y = "Freedom of domestic movement for women", x = "Year")
1584
1585 #       Calculation of country means
1586 free_move_women_countrymeans <- tibble(
1587   countrycode = unique(autocracy_data$cowcode)[order(unique(
1588     autocracy_data$cowcode))],
1589   countrymean = tapply(autocracy_data$v2cldmovem, autocracy_data$cowcode,
1590                       mean) )
1591
1592 #       Histogram and QQ-plot of country means
1593 ggplot(data = free_move_women_countrymeans, aes(x = countrymean)) +

```

```

1594   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1595   labs(x = "Freedom of domestic movement for women", y = "Count") +
1596   blue_light
1597 ggplot(data = free_move_women_countrymeans, aes(sample = countrymean)) +
1598   geom_qq(colour = "lightblue") + geom_qq_line() +
1599   labs(x = "Reference normal distribution",
1600        y = "Freedom of domestic movement for women") + blue_light
1601
1602 #       Compute summary statistics per year
1603 free_move_women_summary <-
1604   tibble(year = sort(unique(autocracy_data$year)),
1605          Mean = tapply(autocracy_data$v2cldmovew, autocracy_data$year, mean),
1606          Q1 = tapply(autocracy_data$v2cldmovew, autocracy_data$year,
1607                    quantile, prob = 0.25),
1608          Median = tapply(autocracy_data$v2cldmovew, autocracy_data$year,
1609                        median),
1610          Q3 = tapply(autocracy_data$v2cldmovew, autocracy_data$year,
1611                    quantile, prob = 0.75),
1612          SD = tapply(autocracy_data$v2cldmovew, autocracy_data$year, sd),
1613          Skew = tapply(autocracy_data$v2cldmovew, autocracy_data$year,
1614                      FUN = Skew),
1615          Kurtosis = tapply(autocracy_data$v2cldmovew, autocracy_data$year,
1616                          FUN = Kurt) )
1617
1618 #       Pivot summary statistics table for use in plotting
1619 free_move_women_summary_longtable <-
1620   pivot_longer(free_move_women_summary, 2:8, names_to = "Statistic",
1621               values_to = "Value")
1622
1623 #       Plot Mean, median, first quartile, third quartile
1624 #       and standard deviation over time
1625 withr::with_options(
1626   list(ggplot2.discrete.colour = lines_palette),
1627   print(ggplot(data = free_move_women_summary_longtable[
1628     free_move_women_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
1629     "Q3", "SD"), ],
1630     aes(x = year, y = Value, colour = Statistic)) +
1631     geom_line() + blue_light +
1632     labs(x = "Year", y = "Freedom of domestic movement for women")) )
1633
1634 #       Plot skew and kurtosis over time
1635 ggplot(data = free_move_women_summary_longtable[
1636   free_move_women_summary_longtable$Statistic %in%
1637   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1638   geom_line() + labs(x = "Year", y = "Freedom of domestic movement for women") +
1639   blue_light
1640
1641 #       Reliability analysis
1642
1643 #       Calculate Cronbach's Alpha by year
1644 free_move_cronbachalphas <-
1645   by(data = autocracy_data[, c("v2cldfmove", "v2cldmovem", "v2cldmovew")],
1646      INDICES = autocracy_data$year, FUN = CronbachAlpha, cond = TRUE,
1647      na.rm = TRUE)
1648
1649 #       Create table for yearly Cronbach's Alpha values
1650 table_free_move_cronbachalphas <-
1651   tibble(year = sort(unique(autocracy_data$year)),
1652          "All items" = NA, "Freedom of foreign movement" = NA,

```

```

1653     "Freedom of domestic movement for men" = NA,
1654     "Freedom of domestic movement for women" = NA)
1655
1656 #   Fill table
1657 for (y in sort(unique(autocracy_data$year))) {
1658   table_free_move_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1659     "All items"] <-
1660     free_move_cronbachalphas[[as.character(y)]]["unconditional"]]
1661   table_free_move_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1662     "Freedom of foreign movement"] <-
1663     free_move_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1664     ]][1, "Cronbach Alpha"]
1665   table_free_move_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1666     "Freedom of domestic movement for men"] <-
1667     free_move_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1668     ]][2, "Cronbach Alpha"]
1669   table_free_move_cronbachalphas[table_free_expr_cronbachalphas$year == y,
1670     "Freedom of domestic movement for women"] <-
1671     free_move_cronbachalphas[[as.character(y)]]["condCronbachAlpha"
1672     ]][3, "Cronbach Alpha"]
1673 }
1674
1675 #   Pivot table for use in plotting
1676 longtab_free_move_cronbachalphas <-
1677   pivot_longer(table_free_move_cronbachalphas, 2:5, names_to = "Type",
1678     values_to = "Value")
1679
1680 #   Plot Cronbach's Alpha by year
1681 ggplot(longtab_free_move_cronbachalphas, aes(x = year, y = Value,
1682     colour = Type)) +
1683   geom_line() + blue_light
1684
1685 #   Calculate range of yearly Cronbach's Alpha values
1686 range(table_free_move_cronbachalphas$"All items")
1687
1688 #   Scale descriptives
1689
1690 #   Entire dataset
1691
1692 #   Heatmap of freedom of movement distribution per year
1693 ggplot(data = autocracy_data, mapping = aes(y = freedom_movement, x = year)) +
1694   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1695   labs(y = "Freedom of movement", x = "Year")
1696
1697 #   Calculation of country means
1698 free_move_full_countrymeans <- tibble(
1699   countrycode = unique(autocracy_data$cowcode)[order(unique(
1700     autocracy_data$cowcode))],
1701   countrymean = tapply(autocracy_data$freedom_movement, autocracy_data$cowcode,
1702     mean)
1703 )
1704 #   Histogram and QQ-plot of country means
1705 ggplot(data = free_move_full_countrymeans, aes(x = countrymean)) +
1706   geom_histogram(binwidth = 0.5, fill = "lightblue") +
1707   labs(x = "Freedom of movement", y = "Count") + blue_light
1708 ggplot(data = free_move_full_countrymeans, aes(sample = countrymean)) +
1709   geom_qq(colour = "lightblue") + geom_qq_line() +
1710   labs(x = "Reference normal distribution", y = "Freedom of movement") +
1711   blue_light

```

```

1712
1713 #       Compute summary statistics per year
1714 free_move_full_summary <-
1715   tibble(year = sort(unique(autocracy_data$year)),
1716         Mean = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1717                       mean),
1718         Q1 = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1719                     quantile, prob = 0.25),
1720         Median = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1721                         median),
1722         Q3 = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1723                     quantile, prob = 0.75),
1724         SD = tapply(autocracy_data$freedom_movement, autocracy_data$year, sd),
1725         Skew = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1726                       FUN = Skew),
1727         Kurtosis = tapply(autocracy_data$freedom_movement, autocracy_data$year,
1728                           FUN = Kurt) )
1729
1730 #       Pivot summary statistics table for use in plotting
1731 free_move_full_summary_longtable <-
1732   pivot_longer(free_move_full_summary, 2:8, names_to = "Statistic",
1733               values_to = "Value")
1734
1735 #       Plot Mean, median, first quartile, third quartile
1736 #       and standard deviation over time
1737 withr::with_options(
1738   list(ggplot2.discrete.colour = lines_palette),
1739   print(ggplot(data = free_move_full_summary_longtable[
1740             free_move_full_summary_longtable$Statistic %in% c("Mean", "Q1",
1741                                                             "Median", "Q3",
1742                                                             "SD"), ],
1743             aes(x = year, y = Value, colour = Statistic)) +
1744             geom_line() + blue_light + labs(x = "Year", y = "Freedom of movement"))) )
1745
1746 #       Plot skew and kurtosis over time
1747 ggplot(data = free_move_full_summary_longtable[
1748       free_move_full_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
1749       aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1750       labs(x = "Year", y = "Freedom of movement") + blue_light
1751
1752 #       Complete cases
1753
1754 #       Heatmap of freedom of movement distribution per year
1755 ggplot(data = datacomplete, mapping = aes(y = freedom_movement, x = year)) +
1756       geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1757       labs(y = "Freedom of movement", x = "Year")
1758
1759 #       Calculation of country means
1760 free_move_complete_countrymeans <- tibble(
1761   countrycode = unique(datacomplete$cowcode)[order(unique(
1762     datacomplete$cowcode))],
1763   countrymean = tapply(datacomplete$freedom_movement, datacomplete$cowcode,
1764                       mean) )
1765
1766 #       Histogram and QQ-plot of country means
1767 ggplot(data = free_move_complete_countrymeans, aes(x = countrymean)) +
1768       geom_histogram(binwidth = 0.5, fill = "lightblue") +
1769       labs(x = "Freedom of movement", y = "Count") + blue_light
1770 ggplot(data = free_move_complete_countrymeans, aes(sample = countrymean)) +

```



```

1771 geom_qq(colour = "lightblue") + geom_qq_line() +
1772 labs(x = "Reference normal distribution", y = "Freedom of movement") +
1773 blue_light
1774
1775 # Compute summary statistics per year
1776 free_move_complete_summary <-
1777 tibble(year = unique(datacomplete$year)[order(unique(
1778 datacomplete$year))],
1779 Mean = tapply(datacomplete$freedom_movement, datacomplete$year, mean),
1780 Q1 = tapply(datacomplete$freedom_movement, datacomplete$year,
1781 quantile, prob = 0.25),
1782 Median = tapply(datacomplete$freedom_movement, datacomplete$year, median),
1783 Q3 = tapply(datacomplete$freedom_movement, datacomplete$year,
1784 quantile, prob = 0.75),
1785 SD = tapply(datacomplete$freedom_movement, datacomplete$year, sd),
1786 Skew = tapply(datacomplete$freedom_movement, datacomplete$year, FUN = Skew),
1787 Kurtosis = tapply(datacomplete$freedom_movement, datacomplete$year,
1788 FUN = Kurt) )
1789
1790 # Pivot summary statistics table for use in plotting
1791 free_move_complete_summary_longtable <-
1792 pivot_longer(free_move_complete_summary, 2:8, names_to = "Statistic",
1793 values_to = "Value")
1794
1795 # Plot Mean, median, first quartile, third quartile
1796 # and standard deviation over time
1797 withr::with_options(
1798 list(ggplot2.discrete.colour = lines_palette),
1799 print(ggplot(data = free_move_complete_summary_longtable[
1800 free_move_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
1801 "Median", "Q3",
1802 "SD"), ],
1803 aes(x = year, y = Value, colour = Statistic)) +
1804 geom_line() + blue_light + labs(x = "Year", y = "Freedom of movement"))) )
1805
1806 # Plot skew and kurtosis over time
1807 ggplot(data = free_move_complete_summary_longtable[
1808 free_move_complete_summary_longtable$Statistic %in%
1809 c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
1810 geom_line() + labs(x = "Year", y = "Freedom of movement") + blue_light
1811
1812 ##### Protection of life and physical integrity descriptives #####
1813
1814 # Item descriptives
1815
1816 # Freedom from torture
1817
1818 # Establish observed range
1819 range(autocracy_data$v2cltort)
1820
1821 # Heatmap of media freedom of discussion for men distribution per year
1822 ggplot(data = autocracy_data, mapping = aes(y = v2cltort, x = year)) +
1823 geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
1824 labs(y = "Freedom from torture", x = "Year")
1825
1826 # Calculation of country means
1827 free_tort_countrymeans <- tibble(
1828 countrycode = unique(autocracy_data$cowcode)[order(unique(
1829 autocracy_data$cowcode))],

```

```

1830     countrymean = tapply(autocracy_data$v2cltort, autocracy_data$cowcode,
1831                           mean)
1832   )
1833   #       Histogram and QQ-plot of country means
1834   ggplot(data = free_tort_countrymeans, aes(x = countrymean)) +
1835     geom_histogram(binwidth = 0.5, fill = "lightblue") +
1836     labs(x = "Freedom from torture", y = "Count") + blue_light
1837   ggplot(data = free_tort_countrymeans, aes(sample = countrymean)) +
1838     geom_qq(colour = "lightblue") + geom_qq_line() +
1839     labs(x = "Reference normal distribution", y = "Freedom from torture") +
1840     blue_light
1841
1842   #       Compute summary statistics per year
1843   free_tort_summary <-
1844     tibble(year = sort(unique(autocracy_data$year)),
1845            Mean = tapply(autocracy_data$v2cltort, autocracy_data$year, mean),
1846            Q1 = tapply(autocracy_data$v2cltort, autocracy_data$year,
1847                       quantile, prob = 0.25),
1848            Median = tapply(autocracy_data$v2cltort, autocracy_data$year, median),
1849            Q3 = tapply(autocracy_data$v2cltort, autocracy_data$year,
1850                       quantile, prob = 0.75),
1851            SD = tapply(autocracy_data$v2cltort, autocracy_data$year, sd),
1852            Skew = tapply(autocracy_data$v2cltort, autocracy_data$year,
1853                         FUN = Skew),
1854            Kurtosis = tapply(autocracy_data$v2cltort, autocracy_data$year,
1855                             FUN = Kurt) )
1856
1857   #       Pivot summary statistics table for use in plotting
1858   free_tort_summary_longtable <-
1859     pivot_longer(free_tort_summary, 2:8, names_to = "Statistic",
1860                 values_to = "Value")
1861
1862   #       Plot Mean, median, first quartile, third quartile
1863   #       and standard deviation over time
1864   withr::with_options(
1865     list(ggplot2.discrete.colour = lines_palette),
1866     print(ggplot(data = free_tort_summary_longtable[
1867               free_tort_summary_longtable$Statistic %in% c("Mean", "Q1",
1868                                                           "Median", "Q3",
1869                                                           "SD"), ],
1870               aes(x = year, y = Value, colour = Statistic)) +
1871               geom_line() + blue_light + labs(x = "Year", y = "Freedom from torture"))) )
1872
1873   #       Plot skew and kurtosis over time
1874   ggplot(data = free_tort_summary_longtable[
1875         free_tort_summary_longtable$Statistic %in%
1876         c("Skew", "Kurtosis"), ],
1877         aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1878         labs(x = "Year", y = "Freedom from torture") + blue_light
1879
1880
1881   #       Freedom from political killings
1882
1883   #       Establish observed range
1884   range(autocracy_data$v2clkill)
1885
1886   #       Heatmap of freedom of discussion for women distribution per year
1887   ggplot(data = autocracy_data, mapping = aes(y = v2clkill, x = year)) +
1888     geom_bin2d(binwidth = c(1, 0.5)) + blue_light +

```

```

1889   labs(y = "Freedom from political killings", x = "Year")
1890
1891   #       Calculation of country means
1892   free_kill_countrymeans <- tibble(
1893     countrycode = unique(autocracy_data$cowcode)[order(unique(
1894       autocracy_data$cowcode))],
1895     countrymean = tapply(autocracy_data$v2clkill, autocracy_data$cowcode,
1896       mean) )
1897   #       Histogram and QQ-plot of country means
1898   ggplot(data = free_kill_countrymeans, aes(x = countrymean)) +
1899     geom_histogram(binwidth = 0.5, fill = "lightblue") +
1900     labs(x = "Freedom from political killings", y = "Count") + blue_light
1901   ggplot(data = free_kill_countrymeans, aes(sample = countrymean)) +
1902     geom_qq(colour = "lightblue") + geom_qq_line() +
1903     labs(x = "Reference normal distribution",
1904       y = "Freedom from political killings") + blue_light
1905
1906   #       Compute summary statistics per year
1907   free_kill_summary <-
1908     tibble(year = sort(unique(autocracy_data$year)),
1909       Mean = tapply(autocracy_data$v2clkill, autocracy_data$year, mean),
1910       Q1 = tapply(autocracy_data$v2clkill, autocracy_data$year,
1911         quantile, prob = 0.25),
1912       Median = tapply(autocracy_data$v2clkill, autocracy_data$year, median),
1913       Q3 = tapply(autocracy_data$v2clkill, autocracy_data$year,
1914         quantile, prob = 0.75),
1915       SD = tapply(autocracy_data$v2clkill, autocracy_data$year, sd),
1916       Skew = tapply(autocracy_data$v2clkill, autocracy_data$year,
1917         FUN = Skew),
1918       Kurtosis = tapply(autocracy_data$v2clkill, autocracy_data$year,
1919         FUN = Kurt) )
1920
1921   #       Pivot summary statistics table for use in plotting
1922   free_kill_summary_longtable <-
1923     pivot_longer(free_kill_summary, 2:8, names_to = "Statistic",
1924       values_to = "Value")
1925
1926   #       Plot Mean, median, first quartile, third quartile
1927   #       and standard deviation over time
1928   withr::with_options(
1929     list(ggplot2.discrete.colour = lines_palette),
1930     print(ggplot(data = free_kill_summary_longtable[
1931       free_kill_summary_longtable$Statistic %in% c("Mean", "Q1", "Median", "Q3",
1932         "SD"), ],
1933       aes(x = year, y = Value, colour = Statistic)) +
1934         geom_line() + blue_light + labs(x = "Year",
1935           y = "Freedom from political killings"))) )
1936
1937   #       Plot skew and kurtosis over time
1938   ggplot(data = free_kill_summary_longtable[
1939     free_kill_summary_longtable$Statistic %in%
1940     c("Skew", "Kurtosis"), ],
1941     aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
1942     labs(x = "Year", y = "Freedom from political killings") + blue_light
1943
1944   #       Item intercorrelation
1945
1946   #       Calculate correlations by year
1947   life_phys_cors <-

```

```

1948   by(data = autocracy_data[, c("v2c1tort", "v2c1kill")],
1949     INDICES = autocracy_data$year, FUN = cor)
1950
1951   #       Create table for correlations
1952   table_life_phys_cors <-
1953     tibble(Year = sort(unique(autocracy_data$year)), Correlation = NA)
1954
1955   #       Fill table
1956   for (y in sort(unique(autocracy_data$year))) {
1957     table_life_phys_cors[table_life_phys_cors$Year == y,
1958       "Correlation"] <- life_phys_cors[[as.character(y)]] [2,1]
1959   }
1960
1961   #       Plot yearly correlations
1962   ggplot(table_life_phys_cors, aes(x = Year, y = Correlation)) +
1963     geom_line(colour = "tomato") + blue_light
1964
1965   #       Establish range for yearly correlations
1966   range(table_life_phys_cors$Correlation)
1967
1968   # Scale descriptives
1969
1970   # Entire dataset
1971
1972   # Establish observed range for scale
1973   range(autocracy_data$life_phys_x100)
1974
1975   # Heatmap of protection of life and physical integrity distribution per year
1976   ggplot(data = autocracy_data, mapping = aes(y = life_phys_x100, x = year)) +
1977     geom_bin2d(binwidth = c(1, 10)) + blue_light +
1978     labs(y = "Protection of life and physical integrity", x = "Year")
1979
1980   # Calculation of country means
1981   full_life_phys_countrymeans <- tibble(
1982     countrycode = unique(autocracy_data$cowcode)[order(unique(
1983       autocracy_data$cowcode))],
1984     countrymean = tapply(autocracy_data$life_phys_x100, autocracy_data$cowcode,
1985       mean)
1986   )
1987   # Histogram and QQ-plot of country means
1988   ggplot(data = full_life_phys_countrymeans, aes(x = countrymean)) +
1989     geom_histogram(binwidth = 10, fill = "lightblue") +
1990     labs(x = "Protection of life and physical integrity", y = "Count") +
1991     blue_light
1992   ggplot(data = full_life_phys_countrymeans, aes(sample = countrymean)) +
1993     geom_qq(colour = "lightblue") + geom_qq_line() +
1994     labs(x = "Reference normal distribution",
1995       y = "Protection of life and physical integrity") + blue_light
1996
1997   # Compute scale summary statistics per year
1998   full_life_phys_summary <-
1999     tibble(year = sort(unique(autocracy_data$year)),
2000       Mean = tapply(autocracy_data$life_phys_x100, autocracy_data$year,
2001         mean),
2002       Q1 = tapply(autocracy_data$life_phys_x100, autocracy_data$year,
2003         quantile, prob = 0.25),
2004       Median = tapply(autocracy_data$life_phys_x100, autocracy_data$year,
2005         median),
2006       Q3 = tapply(autocracy_data$life_phys_x100, autocracy_data$year,

```

```

2007         quantile, prob = 0.75),
2008     SD = tapply(autocracy_data$life_phys_x100, autocracy_data$year, sd),
2009     Skew = tapply(autocracy_data$life_phys_x100, autocracy_data$year,
2010                 FUN = Skew),
2011     Kurtosis = tapply(autocracy_data$life_phys_x100, autocracy_data$year,
2012                     FUN = Kurt) )
2013
2014 #       Pivot summary statistics table for use in plotting
2015 full_life_phys_summary_longtable <-
2016     pivot_longer(full_life_phys_summary, 2:8, names_to = "Statistic",
2017                 values_to = "Value")
2018
2019 #       Plot mean, median, first quartile, third quartile
2020 #       and standard deviation of scale over time
2021 withr::with_options(
2022     list(ggplot2.discrete.colour = lines_palette),
2023     print(ggplot(data = full_life_phys_summary_longtable[
2024         full_life_phys_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
2025             "Q3", "SD"), ],
2026         aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2027         labs(x = "Year", y = "Protection of life and physical integrity")) )
2028
2029 #       Plot skew and kurtosis of scale over time
2030 ggplot(data = full_life_phys_summary_longtable[
2031     full_life_phys_summary_longtable$Statistic %in%
2032     c("Skew", "Kurtosis"), ],
2033     aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2034     labs(x = "Year", y = "Protection of life and physical integrity") + blue_light
2035
2036 #       Complete cases
2037
2038 #       Establish observed range for scale
2039 range(datacomplete$life_phys_x100)
2040
2041 #       Heatmap of protection of life and physical integrity distribution per year
2042 ggplot(data = datacomplete, mapping = aes(y = life_phys_x100, x = year)) +
2043     geom_bin2d(binwidth = c(1, 10)) + blue_light +
2044     labs(y = "Protection of life and physical integrity", x = "Year")
2045
2046 #       Calculation of country means
2047 complete_life_phys_countrymeans <- tibble(
2048     countrycode = unique(datacomplete$cowcode)[order(unique(
2049         datacomplete$cowcode))],
2050     countrymean = tapply(datacomplete$life_phys_x100, datacomplete$cowcode,
2051                         mean)
2052 )
2053 #       Histogram and QQ-plot of country means
2054 ggplot(data = complete_life_phys_countrymeans, aes(x = countrymean)) +
2055     geom_histogram(binwidth = 10, fill = "lightblue") +
2056     labs(x = "Protection of life and physical integrity", y = "Count") +
2057     blue_light
2058 ggplot(data = complete_life_phys_countrymeans, aes(sample = countrymean)) +
2059     geom_qq(colour = "lightblue") + geom_qq_line() +
2060     labs(x = "Reference normal distribution",
2061         y = "Protection of life and physical integrity") + blue_light
2062
2063 #       Compute scale summary statistics per year
2064 complete_life_phys_summary <-
2065     tibble(year = sort(unique(datacomplete$year)),

```

```

2066 Mean = tapply(datacomplete$life_phys_x100, datacomplete$year, mean),
2067 Q1 = tapply(datacomplete$life_phys_x100, datacomplete$year,
2068 quantile, prob = 0.25),
2069 Median = tapply(datacomplete$life_phys_x100, datacomplete$year,
2070 median),
2071 Q3 = tapply(datacomplete$life_phys_x100, datacomplete$year,
2072 quantile, prob = 0.75),
2073 SD = tapply(datacomplete$life_phys_x100, datacomplete$year, sd),
2074 Skew = tapply(datacomplete$life_phys_x100, datacomplete$year,
2075 FUN = Skew),
2076 Kurtosis = tapply(datacomplete$life_phys_x100, datacomplete$year,
2077 FUN = Kurt) )
2078
2079 # Pivot summary statistics table for use in plotting
2080 complete_life_phys_summary_longtable <-
2081 pivot_longer(complete_life_phys_summary, 2:8, names_to = "Statistic",
2082 values_to = "Value")
2083
2084 # Plot mean, median, first quartile, third quartile
2085 # and standard deviation of scale over time
2086 withr::with_options(
2087 list(ggplot2.discrete.colour = lines_palette),
2088 print(ggplot(data = complete_life_phys_summary_longtable[
2089 complete_life_phys_summary_longtable$Statistic %in% c("Mean", "Q1",
2090 "Median", "Q3",
2091 "SD"), ],
2092 aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2093 labs(x = "Year", y = "Protection of life and physical integrity"))) )
2094
2095 # Plot skew and kurtosis of scale over time
2096 ggplot(data = complete_life_phys_summary_longtable[
2097 complete_life_phys_summary_longtable$Statistic %in%
2098 c("Skew", "Kurtosis"), ],
2099 aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2100 labs(x = "Year", y = "Protection of life and physical integrity") + blue_light
2101
2102 ##### Population descriptives #####
2103
2104 # Original variable
2105
2106 # Heatmap of population distribution per year
2107 ggplot(data = autocracy_data, mapping = aes(y = e_mipopula, x = year)) +
2108 geom_bin2d(binwidth = c(1, 10000)) + blue_light +
2109 labs(y = "Population (thousands)", x = "Year")
2110 ggplot(data = autocracy_data[autocracy_data$cowcode == 710, ],
2111 mapping = aes(y = e_mipopula, x = year)) +
2112 geom_bin2d(binwidth = c(1, 10000)) + blue_light +
2113 labs(y = "Population (thousands)", x = "Year")
2114
2115 # Calculation of country means
2116 population_countrymeans <- tibble(
2117 countrycode = unique(autocracy_data$cowcode)[order(unique(
2118 autocracy_data$cowcode))],
2119 countrymean = tapply(autocracy_data$e_mipopula, autocracy_data$cowcode,
2120 mean, na.rm = TRUE)
2121 )
2122 # Histogram and QQ-plot of country means
2123 ggplot(data = population_countrymeans, aes(x = countrymean)) +
2124 geom_histogram(binwidth = 10000, fill = "lightblue") +

```

```

2125   labs(x = "Population (thousands)", y = "Count") + blue_light
2126 ggplot(data = population_countrymeans, aes(sample = countrymean)) +
2127   geom_qq(colour = "lightblue") + geom_qq_line() +
2128   labs(x = "Reference normal distribution", y = "Population (thousands)") +
2129   blue_light
2130
2131 #       Compute summary statistics per year
2132 population_summary <-
2133   tibble(year = sort(unique(autocracy_data$year)),
2134     Mean = tapply(autocracy_data$e_mipopula, autocracy_data$year, mean,
2135       na.rm = TRUE),
2136     Q1 = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2137       quantile, prob = 0.25, na.rm = TRUE),
2138     Median = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2139       median, na.rm = TRUE),
2140     Q3 = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2141       quantile, prob = 0.75, na.rm = TRUE),
2142     SD = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2143       sd, na.rm = TRUE),
2144     Skew = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2145       FUN = Skew, na.rm = TRUE),
2146     Kurtosis = tapply(autocracy_data$e_mipopula, autocracy_data$year,
2147       FUN = Kurt, na.rm = TRUE) )
2148
2149 #       Pivot summary statistics table for use in plotting
2150 population_summary_longtable <-
2151   pivot_longer(population_summary, 2:8, names_to = "Statistic",
2152     values_to = "Value")
2153
2154 #       Plot Mean, median, first quartile, third quartile
2155 #       and standard deviation over time
2156 withr::with_options(
2157   list(ggplot2.discrete.colour = lines_palette),
2158   print(ggplot(data = population_summary_longtable[
2159     population_summary_longtable$Statistic %in% c("Mean", "Q1", "Median", "Q3",
2160       "SD"), ],
2161     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2162     labs(x = "Year", y = "Population (thousands)"))
2163 )
2164
2165 #       Plot skew and kurtosis over time
2166 ggplot(data = population_summary_longtable[
2167   population_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
2168   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2169   labs(x = "Year", y = "Population (thousands)") + blue_light
2170
2171 # Logged variable
2172
2173 # Entire dataset
2174
2175 #       Heatmap of population (log 10) distribution per year
2176 ggplot(data = autocracy_data, mapping = aes(y = log10pop, x = year)) +
2177   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2178   labs(y = "Population (log 10)", x = "Year")
2179
2180 #       Calculation of country means
2181 log_population_full_countrymeans <- tibble(
2182   countrycode = unique(autocracy_data$cowcode)[order(unique(
2183     autocracy_data$cowcode))],

```

```

2184   countrymean = tapply(autocracy_data$log10pop, autocracy_data$cowcode,
2185                        mean, na.rm = TRUE)
2186 )
2187 # Histogram and QQ-plot of country means
2188 ggplot(data = log_population_full_countrymeans, aes(x = countrymean)) +
2189   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2190   labs(x = "Population (log 10)", y = "Count") + blue_light
2191 ggplot(data = log_population_full_countrymeans, aes(sample = countrymean)) +
2192   geom_qq(colour = "lightblue") + geom_qq_line() +
2193   labs(x = "Reference normal distribution", y = "Population (log 10)") +
2194   blue_light
2195
2196 # Compute summary statistics per year
2197 log_population_full_summary <-
2198   tibble(year = sort(unique(autocracy_data$year)),
2199         Mean = tapply(autocracy_data$log10pop, autocracy_data$year, mean,
2200                     na.rm = TRUE),
2201         Q1 = tapply(autocracy_data$log10pop, autocracy_data$year,
2202                   quantile, prob = 0.25, na.rm = TRUE),
2203         Median = tapply(autocracy_data$log10pop, autocracy_data$year, median,
2204                       na.rm = TRUE),
2205         Q3 = tapply(autocracy_data$log10pop, autocracy_data$year,
2206                   quantile, prob = 0.75, na.rm = TRUE),
2207         SD = tapply(autocracy_data$log10pop, autocracy_data$year, sd,
2208                   na.rm = TRUE),
2209         Skew = tapply(autocracy_data$log10pop, autocracy_data$year, FUN = Skew,
2210                      na.rm = TRUE),
2211         Kurtosis = tapply(autocracy_data$log10pop, autocracy_data$year,
2212                          FUN = Kurt, na.rm = TRUE) )
2213
2214 # Pivot summary statistics table for use in plotting
2215 log_population_full_summary_longtable <-
2216   pivot_longer(log_population_full_summary, 2:8, names_to = "Statistic",
2217               values_to = "Value")
2218
2219 # Plot Mean, median, first quartile, third quartile
2220 # and standard deviation over time
2221 withr::with_options(
2222   list(ggplot2.discrete.colour = lines_palette),
2223   print(ggplot(data = log_population_full_summary_longtable[
2224     log_population_full_summary_longtable$Statistic %in% c("Mean", "Q1",
2225                                                         "Median", "Q3",
2226                                                         "SD"), ],
2227             aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2228             labs(x = "Year", y = "Population (log 10)")) )
2229
2230 # Plot skew and kurtosis over time
2231 ggplot(data = log_population_full_summary_longtable[
2232   log_population_full_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
2233   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2234   labs(x = "Year", y = "Population (log 10)") + blue_light
2235
2236 # Complete cases
2237
2238 # Heatmap of population (log 10) distribution per year
2239 ggplot(data = datacomplete, mapping = aes(y = log10pop, x = year)) +
2240   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2241   labs(y = "Population (log 10)", x = "Year")
2242

```



```

2243 # Calculation of country means
2244 log_population_complete_countrymeans <- tibble(
2245   countrycode = unique(datacomplete$cowcode)[order(unique(
2246     datacomplete$cowcode))],
2247   countrymean = tapply(datacomplete$log10pop, datacomplete$cowcode,
2248     mean) )
2249
2250 # Histogram and QQ-plot of country means
2251 ggplot(data = log_population_complete_countrymeans, aes(x = countrymean)) +
2252   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2253   labs(x = "Population (log 10)", y = "Count") + blue_light
2254 ggplot(data = log_population_complete_countrymeans, aes(sample = countrymean)) +
2255   geom_qq(colour = "lightblue") + geom_qq_line() +
2256   labs(x = "Reference normal distribution", y = "Population (log 10)") +
2257   blue_light
2258
2259 # Calculate mean population level
2260 mean(tapply(10^(datacomplete$log10pop), datacomplete$cowcode, mean,
2261   na.rm = TRUE), na.rm = TRUE)
2262
2263 # Compute summary statistics per year
2264 log_population_complete_summary <-
2265   tibble(year = unique(datacomplete$year)[order(unique(
2266     datacomplete$year))],
2267   Mean = tapply(datacomplete$log10pop, datacomplete$year, mean),
2268   Q1 = tapply(datacomplete$log10pop, datacomplete$year,
2269     quantile, prob = 0.25),
2270   Median = tapply(datacomplete$log10pop, datacomplete$year, median),
2271   Q3 = tapply(datacomplete$log10pop, datacomplete$year,
2272     quantile, prob = 0.75),
2273   SD = tapply(datacomplete$log10pop, datacomplete$year, sd),
2274   Skew = tapply(datacomplete$log10pop, datacomplete$year, FUN = Skew),
2275   Kurtosis = tapply(datacomplete$log10pop, datacomplete$year, FUN = Kurt) )
2276
2277 # Pivot summary statistics table for use in plotting
2278 log_population_complete_summary_longtable <-
2279   pivot_longer(log_population_complete_summary, 2:8, names_to = "Statistic",
2280     values_to = "Value")
2281
2282 # Plot Mean, median, first quartile, third quartile
2283 # and standard deviation over time
2284
2285 withr::with_options(
2286   list(ggplot2.discrete.colour = lines_palette),
2287   print(ggplot(data = log_population_complete_summary_longtable[
2288     log_population_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
2289       "Median", "Q3",
2290       "SD"), ],
2291     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2292     labs(x = "Year", y = "Population (log 10)")) )
2293
2294 # Plot skew and kurtosis over time
2295 ggplot(data = log_population_complete_summary_longtable[
2296   log_population_complete_summary_longtable$Statistic %in%
2297   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2298   geom_line() + labs(x = "Year", y = "Population (log 10)") + blue_light
2299
2300 ##### GDP per capita (ln) descriptives #####
2301

```

```

2302 # Entire dataset
2303
2304 # Heatmap of GDP per capita (ln) distribution per year
2305 ggplot(data = autocracy_data, mapping = aes(y = e_migdppcln, x = year)) +
2306   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2307   labs(y = "GDP per capita (ln)", x = "Year")
2308
2309 # Calculation of country means
2310 GDP_cap_full_countrymeans <- tibble(
2311   countrycode = unique(autocracy_data$cowcode)[order(unique(
2312     autocracy_data$cowcode))],
2313   countrymean = tapply(autocracy_data$e_migdppcln, autocracy_data$cowcode,
2314     mean, na.rm = TRUE) )
2315
2316 # Histogram and QQ-plot of country means
2317 ggplot(data = GDP_cap_full_countrymeans, aes(x = countrymean)) +
2318   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2319   labs(x = "GDP per capita (ln)", y = "Count") + blue_light
2320 ggplot(data = GDP_cap_full_countrymeans, aes(sample = countrymean)) +
2321   geom_qq(colour = "lightblue") + geom_qq_line() +
2322   labs(x = "Reference normal distribution", y = "GDP per capita (ln)") +
2323   blue_light
2324
2325 # Compute summary statistics per year
2326 GDP_cap_full_summary <-
2327   tibble(year = sort(unique(autocracy_data$year)),
2328     Mean = tapply(autocracy_data$e_migdppcln, autocracy_data$year, mean,
2329       na.rm = TRUE),
2330     Q1 = tapply(autocracy_data$e_migdppcln, autocracy_data$year,
2331       quantile, prob = 0.25, na.rm = TRUE),
2332     Median = tapply(autocracy_data$e_migdppcln, autocracy_data$year,
2333       median, na.rm = TRUE),
2334     Q3 = tapply(autocracy_data$e_migdppcln, autocracy_data$year,
2335       quantile, prob = 0.75, na.rm = TRUE),
2336     SD = tapply(autocracy_data$e_migdppcln, autocracy_data$year, sd, na.rm =
2337 TRUE),
2338     Skew = tapply(autocracy_data$e_migdppcln, autocracy_data$year,
2339       FUN = Skew, na.rm = TRUE),
2340     Kurtosis = tapply(autocracy_data$e_migdppcln, autocracy_data$year,
2341       FUN = Kurt, na.rm = TRUE) )
2342
2343 # Pivot summary statistics table for use in plotting
2344 GDP_cap_full_summary_longtable <-
2345   pivot_longer(GDP_cap_full_summary, 2:8, names_to = "Statistic",
2346     values_to = "Value")
2347
2348 # Plot Mean, median, first quartile, third quartile
2349 # and standard deviation over time
2350 withr::with_options(
2351   list(ggplot2.discrete.colour = lines_palette),
2352   print(ggplot(data = GDP_cap_full_summary_longtable[
2353     GDP_cap_full_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
2354       "Q3", "SD"), ],
2355     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2356     labs(x = "Year", y = "GDP per capita (ln)")) )
2357
2358 # Plot skew and kurtosis over time
2359 ggplot(data = GDP_cap_full_summary_longtable[
2360   GDP_cap_full_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],

```

```

2361 aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2362 labs(x = "Year", y = "GDP per capita (ln)") + blue_light
2363
2364 # Complete cases
2365
2366 # Heatmap of GDP per capita (ln) distribution per year
2367 ggplot(data = datacomplete, mapping = aes(y = e_migdppcln, x = year)) +
2368 geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2369 labs(y = "GDP per capita (ln)", x = "Year")
2370
2371 # Calculation of country means
2372 GDP_cap_complete_countrymeans <- tibble(
2373   countrycode = unique(datacomplete$cowcode)[order(unique(
2374     datacomplete$cowcode))],
2375   countrymean = tapply(datacomplete$e_migdppcln, datacomplete$cowcode,
2376     mean)
2377 )
2378
2379 # Histogram and QQ-plot of country means
2380 ggplot(data = GDP_cap_complete_countrymeans, aes(x = countrymean)) +
2381 geom_histogram(binwidth = 0.5, fill = "lightblue") +
2382 labs(x = "GDP per capita (ln)", y = "Count") + blue_light
2383 ggplot(data = GDP_cap_complete_countrymeans,
2384   aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
2385 geom_qq_line() + labs(x = "Reference normal distribution",
2386   y = "GDP per capita (ln)") + blue_light
2387
2388 # Calculate mean country mean GDP per capita
2389 mean(tapply(exp(datacomplete$e_migdppcln), datacomplete$cowcode, mean,
2390   na.rm = TRUE), na.rm = TRUE)
2391
2392 # Compute summary statistics per year
2393 GDP_cap_complete_summary <-
2394   tibble(year = unique(datacomplete$year)[order(unique(
2395     datacomplete$year))],
2396   Mean = tapply(datacomplete$e_migdppcln, datacomplete$year, mean),
2397   Q1 = tapply(datacomplete$e_migdppcln, datacomplete$year,
2398     quantile, prob = 0.25),
2399   Median = tapply(datacomplete$e_migdppcln, datacomplete$year, median),
2400   Q3 = tapply(datacomplete$e_migdppcln, datacomplete$year,
2401     quantile, prob = 0.75),
2402   SD = tapply(datacomplete$e_migdppcln, datacomplete$year, sd),
2403   Skew = tapply(datacomplete$e_migdppcln, datacomplete$year, FUN = Skew),
2404   Kurtosis = tapply(datacomplete$e_migdppcln, datacomplete$year, FUN = Kurt) )
2405
2406 # Pivot summary statistics table for use in plotting
2407 GDP_cap_complete_summary_longtable <-
2408   pivot_longer(GDP_cap_complete_summary, 2:8, names_to = "Statistic",
2409     values_to = "Value")
2410
2411 # Plot Mean, median, first quartile, third quartile
2412 # and standard deviation over time
2413
2414 withr::with_options(
2415   list(ggplot2.discrete.colour = lines_palette),
2416   print(ggplot(data = GDP_cap_complete_summary_longtable[
2417     GDP_cap_complete_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
2418       "Q3", "SD"), ],
2419     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +

```

```

2420     labs(x = "Year", y = "GDP per capita (ln)) )
2421
2422 #     Plot skew and kurtosis over time
2423 ggplot(data = GDP_cap_complete_summary_longtable[
2424     GDP_cap_complete_summary_longtable$Statistic %in%
2425     c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2426     geom_line() + labs(x = "Year", y = "GDP per capita (ln)") + blue_light
2427
2428 ##### GDP growth descriptives #####
2429
2430 # Entire dataset
2431
2432 # Heatmap of GDP growth distribution per year
2433 ggplot(data = autocracy_data, mapping = aes(y = e_migdpgro, x = year)) +
2434     geom_bin2d(binwidth = c(1, 0.01)) + blue_light +
2435     labs(y = "GDP growth", x = "Year")
2436
2437 # View outliers
2438 autocracy_data[abs(autocracy_data$e_migdpgro) > 0.5 &
2439     is.na(autocracy_data$e_migdpgro) == FALSE,
2440     c("cowcode", "year", "gwf_country", "e_migdpgro")]
2441
2442 # Calculation of country means
2443 GDP_growth_full_countrymeans <- tibble(
2444     countrycode = unique(autocracy_data$cowcode)[order(unique(
2445     autocracy_data$cowcode))],
2446     countrymean = tapply(autocracy_data$e_migdpgro, autocracy_data$cowcode,
2447     mean, na.rm = TRUE) )
2448
2449 # Histogram and QQ-plot of country means
2450 ggplot(data = GDP_growth_full_countrymeans, aes(x = countrymean)) +
2451     geom_histogram(binwidth = 0.01, fill = "lightblue") +
2452     labs(x = "GDP growth", y = "Count") + blue_light
2453 ggplot(data = GDP_growth_full_countrymeans, aes(sample = countrymean)) +
2454     geom_qq(colour = "lightblue") + geom_qq_line() +
2455     labs(x = "Reference normal distribution", y = "GDP growth") + blue_light
2456
2457 # Compute summary statistics per year
2458 GDP_growth_full_summary <-
2459     tibble(year = sort(unique(autocracy_data$year)),
2460     Mean = tapply(autocracy_data$e_migdpgro, autocracy_data$year, mean,
2461     na.rm = TRUE),
2462     Q1 = tapply(autocracy_data$e_migdpgro, autocracy_data$year,
2463     quantile, prob = 0.25, na.rm = TRUE),
2464     Median = tapply(autocracy_data$e_migdpgro, autocracy_data$year, median,
2465     na.rm = TRUE),
2466     Q3 = tapply(autocracy_data$e_migdpgro, autocracy_data$year,
2467     quantile, prob = 0.75, na.rm = TRUE),
2468     SD = tapply(autocracy_data$e_migdpgro, autocracy_data$year, sd,
2469     na.rm = TRUE),
2470     Skew = tapply(autocracy_data$e_migdpgro, autocracy_data$year,
2471     FUN = Skew, na.rm = TRUE),
2472     Kurtosis = tapply(autocracy_data$e_migdpgro, autocracy_data$year,
2473     FUN = Kurt, na.rm = TRUE) )
2474
2475 # Pivot summary statistics table for use in plotting
2476 GDP_growth_full_summary_longtable <-
2477     pivot_longer(GDP_growth_full_summary, 2:8, names_to = "Statistic",
2478     values_to = "Value")

```

```

2479
2480 # Plot Mean, median, first quartile, third quartile and standard deviation
2481 # over time
2482 withr::with_options(
2483   list(ggplot2.discrete.colour = lines_palette),
2484   print(ggplot(data = GDP_growth_full_summary_longtable[
2485     GDP_growth_full_summary_longtable$Statistic %in% c("Mean", "Q1", "Median",
2486       "Q3", "SD"), ],
2487     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2488     labs(x = "Year", y = "GDP growth")) )
2489
2490 # Plot skew and kurtosis over time
2491 ggplot(data = GDP_growth_full_summary_longtable[
2492   GDP_growth_full_summary_longtable$Statistic %in% c("Skew", "Kurtosis"), ],
2493   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2494   labs(x = "Year", y = "GDP growth") + blue_light
2495
2496 # Complete cases
2497
2498 # Heatmap of GDP growth distribution per year
2499 ggplot(data = datacomplete, mapping = aes(y = e_migdpgro, x = year)) +
2500   geom_bin2d(binwidth = c(1, 0.01)) + blue_light +
2501   labs(y = "GDP growth", x = "Year")
2502
2503 # See which outliers remain
2504 datacomplete[abs(datacomplete$e_migdpgro) > 0.5 &
2505   is.na(datacomplete$e_migdpgro) == FALSE,
2506   c("cowcode", "year", "gwf_country", "e_migdpgro")]
2507
2508
2509 # Calculation of country means
2510 GDP_growth_complete_countrymeans <- tibble(
2511   countrycode = unique(datacomplete$cowcode)[order(unique(
2512     datacomplete$cowcode))],
2513   countrymean = tapply(datacomplete$e_migdpgro, datacomplete$cowcode,
2514     mean) )
2515
2516 # Histogram and QQ-plot of country means
2517 ggplot(data = GDP_growth_complete_countrymeans, aes(x = countrymean)) +
2518   geom_histogram(binwidth = 0.01, fill = "lightblue") +
2519   labs(x = "GDP growth", y = "Count") + blue_light
2520 ggplot(data = GDP_growth_complete_countrymeans, aes(sample = countrymean)) +
2521   geom_qq(colour = "lightblue") + geom_qq_line() +
2522   labs(x = "Reference normal distribution", y = "GDP growth") + blue_light
2523
2524 # Compute summary statistics per year
2525 GDP_growth_complete_summary <-
2526   tibble(year = unique(datacomplete$year)[order(unique(
2527     datacomplete$year))],
2528     Mean = tapply(datacomplete$e_migdpgro, datacomplete$year, mean),
2529     Q1 = tapply(datacomplete$e_migdpgro, datacomplete$year,
2530       quantile, prob = 0.25),
2531     Median = tapply(datacomplete$e_migdpgro, datacomplete$year, median),
2532     Q3 = tapply(datacomplete$e_migdpgro, datacomplete$year,
2533       quantile, prob = 0.75),
2534     SD = tapply(datacomplete$e_migdpgro, datacomplete$year, sd),
2535     Skew = tapply(datacomplete$e_migdpgro, datacomplete$year, FUN = Skew),
2536     Kurtosis = tapply(datacomplete$e_migdpgro, datacomplete$year, FUN = Kurt) )
2537

```

```

2538 #       Pivot summary statistics table for use in plotting
2539 GDP_growth_complete_summary_longtable <-
2540   pivot_longer(GDP_growth_complete_summary, 2:8, names_to = "Statistic",
2541               values_to = "Value")
2542
2543 #   Plot Mean, median, first quartile, third quartile and standard deviation
2544 #   over time
2545 withr::with_options(
2546   list(ggplot2.discrete.colour = lines_palette),
2547   print(ggplot(data = GDP_growth_complete_summary_longtable[
2548     GDP_growth_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
2549     "Median", "Q3",
2550     "SD"), ],
2551     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2552     labs(x = "Year", y = "GDP growth")) )
2553
2554 #   Plot skew and kurtosis over time
2555 ggplot(data = GDP_growth_complete_summary_longtable[
2556   GDP_growth_complete_summary_longtable$Statistic %in%
2557   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2558   geom_line() + labs(x = "Year", y = "GDP growth") + blue_light
2559
2560 ##### Regime type descriptives #####
2561
2562 # Check for coding errors in regime type
2563 table(list(autocracy_data$gwf_party, autocracy_data$gwf_personal,
2564           autocracy_data$gwf_military, autocracy_data$gwf_monarch))
2565
2566 # Create a single regime type factor
2567 autocracy_data$regime_type[autocracy_data$gwf_party == 1] <- "Single-party"
2568 autocracy_data$regime_type[autocracy_data$gwf_personal == 1] <- "Personalist"
2569 autocracy_data$regime_type[autocracy_data$gwf_military == 1] <- "Military"
2570 autocracy_data$regime_type[autocracy_data$gwf_monarch == 1] <- "Monarch"
2571 datacomplete$regime_type[datacomplete$gwf_party == 1] <- "Single-party"
2572 datacomplete$regime_type[datacomplete$gwf_personal == 1] <- "Personalist"
2573 datacomplete$regime_type[datacomplete$gwf_military == 1] <- "Military"
2574 datacomplete$regime_type[datacomplete$gwf_monarch == 1] <- "Monarch"
2575
2576 # Entire dataset
2577
2578 #       Table and chi-square test
2579 full_regime_type_table <- table(autocracy_data$regime_type, autocracy_data$year)
2580 full_regime_type_table
2581 summary(full_regime_type_table)
2582 chisq.test(autocracy_data$regime_type, autocracy_data$year)$stdres
2583 chisq.test(autocracy_data$regime_type, autocracy_data$year)$expected
2584 fisher.test(autocracy_data$regime_type, autocracy_data$year,
2585            simulate.p.value = TRUE, B = 10000)
2586
2587 #       Stacked barplot per year
2588 ggplot(autocracy_data, aes(x = year, fill = regime_type)) +
2589   geom_bar(position = "fill") + blue_light +
2590   labs(x = "Year", y = "Proportion", fill = "Regime type")
2591
2592 # Complete cases
2593
2594 #       Table and chi-square test
2595 complete_regime_type_table <- table(datacomplete$regime_type, datacomplete$year)
2596 complete_regime_type_table

```

```

2597 summary(complete_regime_type_table)
2598 chisq.test(datacomplete$regime_type, datacomplete$year)$expected
2599 fisher.test(datacomplete$regime_type, datacomplete$year,
2600             simulate.p.value = TRUE, B = 10000)
2601
2602 # Stacked barplot per year
2603 ggplot(datacomplete, aes(x = year, fill = regime_type)) +
2604   geom_bar(position = "fill") + blue_light +
2605   labs(x = "Year", y = "Proportion", fill = "Regime type")
2606
2607 ##### International conflict descriptives #####
2608
2609 # Entire dataset
2610
2611 # Table and chi-square test
2612 full_international_conflict_table <- table(autocracy_data$e_miinteco,
2613                                           autocracy_data$year)
2614 full_international_conflict_table
2615 chisq.test(autocracy_data$e_miinteco, autocracy_data$year)
2616 chisq.test(autocracy_data$e_miinteco, autocracy_data$year)$expected
2617 fisher.test(autocracy_data$e_miinteco, autocracy_data$year,
2618            simulate.p.value = TRUE, B = 10000)
2619
2620 # Stacked barplot per year
2621 ggplot(autocracy_data, aes(x = year, fill = as.factor(e_miinteco))) +
2622   geom_bar(position = "fill") + blue_light +
2623   labs(x = "Year", y = "Proportion", fill = "International conflict") +
2624   scale_fill_manual(values = c("lightblue", "darkgreen"),
2625                    labels = c("No international conflict",
2626                              "International conflict"))
2627
2628 # Complete cases
2629
2630 # Table and chi-square test
2631 complete_international_conflict_table <- table(datacomplete$e_miinteco,
2632                                               datacomplete$year)
2633 complete_international_conflict_table
2634 summary(complete_international_conflict_table)
2635 chisq.test(datacomplete$e_miinteco, datacomplete$year)$expected
2636 fisher.test(datacomplete$e_miinteco, datacomplete$year,
2637            simulate.p.value = TRUE, B = 10000)
2638
2639 # Stacked barplot per year
2640 ggplot(datacomplete, aes(x = year, fill = as.factor(e_miinteco))) +
2641   geom_bar(position = "fill") + blue_light +
2642   labs(x = "Year", y = "Proportion", fill = "International conflict") +
2643   scale_fill_manual(values = c("lightblue", "darkgreen"),
2644                    labels = c("No international conflict",
2645                              "International conflict"))
2646
2647 ##### Internal conflict descriptives #####
2648
2649 # Entire dataset
2650
2651 # Table and chi-square test
2652 full_internal_conflict_table <- table(autocracy_data$e_miinterc,
2653                                     autocracy_data$year)
2654 full_internal_conflict_table
2655 chisq.test(autocracy_data$e_miinterc, autocracy_data$year)

```

```

2656 chisq.test(autocracy_data$e_miinterc, autocracy_data$year)$expected
2657
2658 # Stacked barplot per year
2659 ggplot(autocracy_data, aes(x = year, fill = as.factor(e_miinterc))) +
2660   geom_bar(position = "fill") + blue_light +
2661   labs(x = "Year", y = "Proportion", fill = "Internal conflict") +
2662   scale_fill_manual(values = c("lightblue", "darkgreen"),
2663     labels = c("No internal conflict",
2664       "Internal conflict"))
2665
2666 # Complete cases
2667
2668 # Table and chi-square test
2669 complete_internal_conflict_table <- table(datacomplete$e_miinterc,
2670   datacomplete$year)
2671 complete_internal_conflict_table
2672 summary(complete_internal_conflict_table)
2673 chisq.test(datacomplete$e_miinterc, datacomplete$year)$expected
2674 fisher.test(autocracy_data$e_miinterc, autocracy_data$year,
2675   simulate.p.value = TRUE, B = 10000)
2676
2677 # Stacked barplot per year
2678 ggplot(datacomplete, aes(x = year, fill = as.factor(e_miinterc))) +
2679   geom_bar(position = "fill") + blue_light +
2680   labs(x = "Year", y = "Proportion", fill = "Internal conflict") +
2681   scale_fill_manual(values = c("lightblue", "darkgreen"),
2682     labels = c("No internal conflict", "Internal conflict"))
2683
2684 ##### Political violence descriptives #####
2685
2686 # Entire dataset
2687
2688 # Heatmap of political violence distribution per year
2689 ggplot(data = autocracy_data, mapping = aes(y = v2caviol, x = year)) +
2690   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2691   labs(y = "Political violence", x = "Year")
2692
2693 # Calculation of country means
2694 political_violence_full_countrymeans <- tibble(
2695   countrycode = unique(autocracy_data$cowcode)[order(unique(
2696     autocracy_data$cowcode))],
2697   countrymean = tapply(autocracy_data$v2caviol, autocracy_data$cowcode,
2698     mean, na.rm = TRUE)
2699 )
2700 # Histogram and QQ-plot of country means
2701 ggplot(data = political_violence_full_countrymeans, aes(x = countrymean)) +
2702   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2703   labs(x = "Political violence", y = "Count") + blue_light
2704 ggplot(data = political_violence_full_countrymeans, aes(sample = countrymean)) +
2705   geom_qq(colour = "lightblue") + geom_qq_line() +
2706   labs(x = "Reference normal distribution", y = "Political violence") +
2707   blue_light
2708
2709 # Compute summary statistics per year
2710 political_violence_full_summary <-
2711   tibble(year = sort(unique(autocracy_data$year)),
2712     Mean = tapply(autocracy_data$v2caviol, autocracy_data$year, mean,
2713       na.rm = TRUE),
2714     Q1 = tapply(autocracy_data$v2caviol, autocracy_data$year,

```



```

2715         quantile, prob = 0.25, na.rm = TRUE),
2716 Median = tapply(autocracy_data$v2caviol, autocracy_data$year, median,
2717                 na.rm = TRUE),
2718 Q3 = tapply(autocracy_data$v2caviol, autocracy_data$year,
2719            quantile, prob = 0.75, na.rm = TRUE),
2720 SD = tapply(autocracy_data$v2caviol, autocracy_data$year, sd,
2721            na.rm = TRUE),
2722 Skew = tapply(autocracy_data$v2caviol, autocracy_data$year, FUN = Skew,
2723             na.rm = TRUE),
2724 Kurtosis = tapply(autocracy_data$v2caviol, autocracy_data$year,
2725                 FUN = Kurt, na.rm = TRUE) )
2726
2727 # Pivot summary statistics table for use in plotting
2728 political_violence_full_summary_longtable <-
2729   pivot_longer(political_violence_full_summary, 2:8, names_to = "Statistic",
2730               values_to = "Value")
2731
2732 # Plot Mean, median, first quartile, third quartile
2733 # and standard deviation over time
2734 withr::with_options(
2735   list(ggplot2.discrete.colour = lines_palette),
2736   print(ggplot(data = political_violence_full_summary_longtable[
2737             political_violence_full_summary_longtable$Statistic %in% c("Mean", "Q1",
2738                               "Median", "Q3",
2739                               "SD"), ],
2740             aes(x = year, y = Value, colour = Statistic)) +
2741             geom_line() + blue_light + labs(x = "Year", y = "Political violence"))) )
2742
2743 # Plot skew and kurtosis over time
2744 ggplot(data = political_violence_full_summary_longtable[
2745   political_violence_full_summary_longtable$Statistic %in%
2746   c("Skew", "Kurtosis"), ],
2747   aes(x = year, y = Value, linetype = Statistic)) + geom_line() +
2748   labs(x = "Year", y = "Political violence") + blue_light
2749
2750 # Complete cases
2751
2752 # Heatmap of political violence distribution per year
2753 ggplot(data = datacomplete, mapping = aes(y = v2caviol, x = year)) +
2754   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2755   labs(y = "Political violence", x = "Year")
2756
2757 # Calculation of country means
2758 political_violence_complete_countrymeans <- tibble(
2759   countrycode = unique(datacomplete$cowcode)[order(unique(
2760     datacomplete$cowcode))],
2761   countrymean = tapply(datacomplete$v2caviol, datacomplete$cowcode, mean) )
2762
2763 # Histogram and QQ-plot of country means
2764 ggplot(data = political_violence_complete_countrymeans, aes(x = countrymean)) +
2765   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2766   labs(x = "Political violence", y = "Count") + blue_light
2767 ggplot(data = political_violence_complete_countrymeans,
2768   aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
2769   geom_qq_line() + labs(x = "Reference normal distribution",
2770   y = "Political violence") + blue_light
2771
2772 # Compute summary statistics per year
2773 political_violence_complete_summary <-

```

```

2774   tibble(year = unique(datacomplete$year)[order(unique(datacomplete$year))],
2775         Mean = tapply(datacomplete$v2caviol, datacomplete$year, mean),
2776         Q1 = tapply(datacomplete$v2caviol, datacomplete$year,
2777                   quantile, prob = 0.25),
2778         Median = tapply(datacomplete$v2caviol, datacomplete$year, median),
2779         Q3 = tapply(datacomplete$v2caviol, datacomplete$year,
2780                   quantile, prob = 0.75),
2781         SD = tapply(datacomplete$v2caviol, datacomplete$year, sd),
2782         Skew = tapply(datacomplete$v2caviol, datacomplete$year, FUN = Skew),
2783         Kurtosis = tapply(datacomplete$v2caviol, datacomplete$year, FUN = Kurt) )
2784
2785   #       Pivot summary statistics table for use in plotting
2786   political_violence_complete_summary_longtable <-
2787     pivot_longer(political_violence_complete_summary, 2:8, names_to = "Statistic",
2788                 values_to = "Value")
2789
2790   #       Plot Mean, median, first quartile, third quartile
2791   #       and standard deviation over time
2792   withr::with_options(
2793     list(ggplot2.discrete.colour = lines_palette),
2794     print(ggplot(data = political_violence_complete_summary_longtable[
2795               political_violence_complete_summary_longtable$Statistic %in% c("Mean", "Q1",
2796                                     "Median",
2797                                     "Q3", "SD"),
2798             ], aes(x = year, y = Value, colour = Statistic)) + geom_line() +
2799               blue_light + labs(x = "Year", y = "Political violence"))) )
2800
2801   #       Plot skew and kurtosis over time
2802   ggplot(data = political_violence_complete_summary_longtable[
2803         political_violence_complete_summary_longtable$Statistic %in%
2804         c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2805     geom_line() + labs(x = "Year", y = "Political violence") + blue_light
2806
2807   ##### Rigour and impartiality public administration descriptives #####
2808
2809   # Entire dataset
2810
2811   #       Heatmap of political violence distribution per year
2812   ggplot(data = autocracy_data, mapping = aes(y = v2clrspct, x = year)) +
2813     geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2814     labs(y = "Rig. & impart. pub. admin.", x = "Year")
2815
2816   #       Calculation of country means
2817   rigour_impartiality_full_countrymeans <- tibble(
2818     countrycode = unique(autocracy_data$cowcode)[order(unique(
2819       autocracy_data$cowcode))],
2820     countrymean = tapply(autocracy_data$v2clrspct, autocracy_data$cowcode, mean) )
2821
2822   #       Histogram and QQ-plot of country means
2823   ggplot(data = rigour_impartiality_full_countrymeans, aes(x = countrymean)) +
2824     geom_histogram(binwidth = 0.5, fill = "lightblue") +
2825     labs(x = "Rig. & impart. pub. admin.", y = "Count") + blue_light
2826   ggplot(data = rigour_impartiality_full_countrymeans, aes(sample = countrymean)) +
2827     geom_qq(colour = "lightblue") + geom_qq_line() +
2828     labs(x = "Reference normal distribution", y = "Rig. & impart. pub. admin.") +
2829     blue_light
2830
2831   #       Compute summary statistics per year
2832   rigour_impartiality_full_summary <-

```

```

2833   tibble(year = sort(unique(autocracy_data$year)),
2834           Mean = tapply(autocracy_data$v2clrspct, autocracy_data$year, mean),
2835           Q1 = tapply(autocracy_data$v2clrspct, autocracy_data$year,
2836                       quantile, prob = 0.25),
2837           Median = tapply(autocracy_data$v2clrspct, autocracy_data$year, median),
2838           Q3 = tapply(autocracy_data$v2clrspct, autocracy_data$year,
2839                       quantile, prob = 0.75),
2840           SD = tapply(autocracy_data$v2clrspct, autocracy_data$year, sd),
2841           Skew = tapply(autocracy_data$v2clrspct, autocracy_data$year,
2842                         FUN = Skew),
2843           Kurtosis = tapply(autocracy_data$v2clrspct, autocracy_data$year,
2844                             FUN = Kurt) )
2845
2846 #       Pivot summary statistics table for use in plotting
2847 rigour_impartiality_full_summary_longtable <-
2848   pivot_longer(rigour_impartiality_full_summary, 2:8, names_to = "Statistic",
2849               values_to = "Value")
2850
2851 #       Plot Mean, median, first quartile, third quartile
2852 #       and standard deviation over time
2853 withr::with_options(
2854   list(ggplot2.discrete.colour = lines_palette),
2855   print(ggplot(data = rigour_impartiality_full_summary_longtable[
2856               rigour_impartiality_full_summary_longtable$Statistic %in% c("Mean", "Q1",
2857                                       "Median", "Q3",
2858                                       "SD"), ],
2859             aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2860             labs(x = "Year", y = "Rig. & impart. pub. admin.)) )
2861
2862 #       Plot skew and kurtosis over time
2863 ggplot(data = rigour_impartiality_full_summary_longtable[
2864       rigour_impartiality_full_summary_longtable$Statistic %in%
2865       c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2866   geom_line() + labs(x = "Year", y = "Rig. & impart. pub. admin.") + blue_light
2867
2868 #       Complete cases
2869
2870 #       Heatmap of political violence distribution per year
2871 ggplot(data = datacomplete, mapping = aes(y = v2clrspct, x = year)) +
2872   geom_bin2d(binwidth = c(1, 0.5)) + blue_light +
2873   labs(y = "Rig. & impart. pub. admin.", x = "Year")
2874
2875 #       Calculation of country means
2876 rigour_impartiality_complete_countrymeans <- tibble(
2877   countrycode = unique(datacomplete$cowcode)[order(unique(
2878     datacomplete$cowcode))],
2879   countrymean = tapply(datacomplete$v2clrspct, datacomplete$cowcode, mean) )
2880
2881 #       Histogram and QQ-plot of country means
2882 ggplot(data = rigour_impartiality_complete_countrymeans, aes(x = countrymean)) +
2883   geom_histogram(binwidth = 0.5, fill = "lightblue") +
2884   labs(x = "Rig. & impart. pub. admin.", y = "Count") + blue_light
2885 ggplot(data = rigour_impartiality_complete_countrymeans,
2886       aes(sample = countrymean)) + geom_qq(colour = "lightblue") +
2887   geom_qq_line() + labs(x = "Reference normal distribution",
2888                       y = "Rig. & impart. pub. admin.") + blue_light
2889
2890 #       Compute summary statistics per year
2891 rigour_impartiality_complete_summary <-

```

```

2892   tibble(year = unique(datacomplete$year)[order(unique(datacomplete$year))],
2893     Mean = tapply(datacomplete$v2clrspct, datacomplete$year, mean),
2894     Q1 = tapply(datacomplete$v2clrspct, datacomplete$year,
2895       quantile, prob = 0.25),
2896     Median = tapply(datacomplete$v2clrspct, datacomplete$year, median),
2897     Q3 = tapply(datacomplete$v2clrspct, datacomplete$year,
2898       quantile, prob = 0.75),
2899     SD = tapply(datacomplete$v2clrspct, datacomplete$year, sd),
2900     Skew = tapply(datacomplete$v2clrspct, datacomplete$year, FUN = Skew),
2901     Kurtosis = tapply(datacomplete$v2clrspct, datacomplete$year, FUN = Kurt) )
2902
2903 #       Pivot summary statistics table for use in plotting
2904 rigour_impartiality_complete_summary_longtable <-
2905   pivot_longer(rigour_impartiality_complete_summary, 2:8,
2906     names_to = "Statistic", values_to = "Value")
2907
2908 #       Plot Mean, median, first quartile, third quartile
2909 #       and standard deviation over time
2910 withr::with_options(
2911   list(ggplot2.discrete.colour = lines_palette),
2912   print(ggplot(data = rigour_impartiality_complete_summary_longtable[
2913     rigour_impartiality_complete_summary_longtable$Statistic %in%
2914     c("Mean", "Q1", "Median", "Q3", "SD"), ],
2915     aes(x = year, y = Value, colour = Statistic)) + geom_line() + blue_light +
2916     labs(x = "Year", y = "Rig. & impart. pub. admin.")) )
2917
2918 #       Plot skew and kurtosis over time
2919 ggplot(data = rigour_impartiality_complete_summary_longtable[
2920   rigour_impartiality_complete_summary_longtable$Statistic %in%
2921   c("Skew", "Kurtosis"), ], aes(x = year, y = Value, linetype = Statistic)) +
2922   geom_line() + labs(x = "Year", y = "Rig. & impart. pub. admin.") + blue_light
2923
2924 ##### Missing data analysis #####
2925
2926 # Remove variables that are uninformative during missing data analysis
2927 autocracy_data_clean <- autocracy_data[, c("cowcode", "gwf_country", "year",
2928   "latent_personalism", "regime_type",
2929   "v2x_clphy", "free_expr_x100",
2930   "e_miinteco", "e_miinterc",
2931   "e_migdppln", "e_migdpgro",
2932   "v2caviol", "v2clrspct", "log10pop",
2933   "v2caassemb",
2934   "freedom_movement")]
2935 datacomplete_clean <- datacomplete[, c("cowcode", "gwf_country", "year",
2936   "latent_personalism", "regime_type",
2937   "v2x_clphy", "free_expr_x100",
2938   "e_miinteco", "e_miinterc",
2939   "e_migdppln", "e_migdpgro",
2940   "v2caviol", "v2clrspct", "log10pop",
2941   "v2caassemb", "freedom_movement")]
2942
2943 # Tabulate the distribution of values over years and cowcodes
2944 table(autocracy_data$year)
2945 table(autocracy_data$cowcode)
2946
2947 # Single out countries observed in all years
2948 table(autocracy_data$cowcode)[table(autocracy_data$cowcode) == 65]
2949
2950 # Visualise the distribution of values over years and cowcodes

```

```

2951 ggplot(autocracy_data_clean, aes(x = year)) + geom_bar(fill = "lightblue") +
2952   labs(x = "Year", y = "Count") + blue_light
2953 ggplot(autocracy_data_clean, aes(x = as.factor(cowcode))) +
2954   geom_bar(fill = "lightblue") + labs(x = "Country code", y = "Count") +
2955   guides(x = guide_axis(angle = 90)) + blue_light
2956
2957 # Table of missing values per variable
2958 missing_value_table <- tibble(Variable =
2959   variable.names(autocracy_data_clean)[4:16],
2960   N = NA, "N%" = NA, n = NA, "n%" = NA, T_min = NA,
2961   T_max = NA)
2962
2963 for(V in variable.names(autocracy_data_clean)[4:16]) {
2964   missing_value_table[missing_value_table$Variable == V, "N"] <-
2965     nrow(autocracy_data[complete.cases(autocracy_data[, c("cowcode", "year",
2966       V])], ])
2967   missing_value_table[missing_value_table$Variable == V, "T_min"] <-
2968     min(table(autocracy_data$cowcode[complete.cases(autocracy_data[, V])]))
2969   missing_value_table[missing_value_table$Variable == V, "T_max"] <-
2970     max(table(autocracy_data$cowcode[complete.cases(autocracy_data[, V])]))
2971   missing_value_table[missing_value_table$Variable == V, "n"] <-
2972     length(unique(autocracy_data$cowcode[
2973       complete.cases(autocracy_data[, c("cowcode", "year", V)])]))
2974 }
2975 missing_value_table$`N%` <- (missing_value_table$N / 4591) * 100
2976 missing_value_table$`n%` <- (missing_value_table$n / 119) * 100
2977
2978 # Create a variable listing which variables a case misses
2979 autocracy_data_clean$which_miss <- "Missing:"
2980
2981 for (V in variable.names(autocracy_data_clean)) {
2982   autocracy_data_clean[is.na(autocracy_data_clean[, V]),"which_miss"] <-
2983     paste(autocracy_data_clean[is.na(autocracy_data_clean[, V]),"which_miss"],
2984       V, sep = " ")
2985 }
2986
2987 # Visualise missingness over countries and time
2988 ggplot(autocracy_data_clean, aes(x = year, y = as.factor(cowcode),
2989   fill = which_miss)) +
2990   labs(x = "Year", y = "Country code") +
2991   scale_fill_discrete(name = "Missing",
2992     labels = c("None", "GDP growth",
2993       "GDP growth &\nPopulation (log 10)",
2994       "GDP growth,\nPolitical violence, \nFreedom of
2995 assembly",
2996       "GDP per capita (ln)",
2997       "GDP per capita (ln) &\nGDP growth",
2998       "GDP per capita (ln),\nGDP growth, &\nPopulation
2999 (log 10)",
3000       "GDP per capita (ln),\nGDP growth,\nPolitical
3001 violence,\nFreedom of assembly",
3002       "Internal & International conflict", "GDP
3003 growth,\nInternal, & International conflict",
3004       "GDP per capita (ln),\nGDP growth,\nInternal &
3005 International conflict",
3006       "GDP per capita (ln),\nGDP growth,\nPopulation
3007 (log 10),\nInternal & International conflict",

```

```

3008         "GDP per capita (ln),\nGDP growth,\nPolitical
3009 violence,\nPopulation (log 10),\nFreedom of assembly,\nInternal & International
3010 conflict",
3011         "Population (log 10),\nInternal & International
3012 conflict",
3013         "Political violence,\nPopulation (log
3014 10),\nFreedom of assembly,\nInternal & International conflict",
3015         "Population (log 10)",
3016         "Political violence",
3017         "Political violence &\nFreedom of assembly") ) +
3018     guides(x = guide_axis(angle = 90)) + blue_light + geom_tile()
3019
3020 # Compare distributions in full dataset and subset of complete cases
3021
3022 # Summary statistics table
3023
3024 # Create table
3025 comparison_table <- tibble(Variable = rep(c("latent_personalism",
3026     "life_phys_x100",
3027     "free_expr_x100", "e_migdppcIn",
3028     "e_migdpgro", "v2caviol",
3029     "v2clrspct", "log10pop",
3030     "freedom_movement",
3031     "v2caassemb"), each = 2),
3032     Dataset = rep(c("Full", "Complete"), times = 10),
3033     Mean = NA, SD = NA, Minimum = NA, Maximum = NA,
3034     Skew = NA, Kurt = NA, ICC = NA)
3035
3036 # Fill table
3037 for (V in c("latent_personalism", "life_phys_x100", "free_expr_x100",
3038     "e_migdppcIn", "e_migdpgro", "v2caviol", "v2clrspct", "log10pop",
3039     "freedom_movement", "v2caassemb")) {
3040     comparison_table[comparison_table$Variable == V &
3041         comparison_table$Dataset == "Full", "Mean"] <-
3042         mean(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,
3043             na.rm = TRUE), na.rm = TRUE)
3044     comparison_table[comparison_table$Variable == V &
3045         comparison_table$Dataset == "Complete", "Mean"] <-
3046         mean(tapply(datacomplete[, V], datacomplete$cowcode, mean,
3047             na.rm = TRUE), na.rm = TRUE)
3048     comparison_table[comparison_table$Variable == V &
3049         comparison_table$Dataset == "Full", "SD"] <-
3050         sd(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,
3051             na.rm = TRUE), na.rm = TRUE)
3052     comparison_table[comparison_table$Variable == V &
3053         comparison_table$Dataset == "Complete", "SD"] <-
3054         sd(tapply(datacomplete[, V], datacomplete$cowcode, mean,
3055             na.rm = TRUE), na.rm = TRUE)
3056     comparison_table[comparison_table$Variable == V &
3057         comparison_table$Dataset == "Full", "Minimum"] <-
3058         min(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,
3059             na.rm = TRUE), na.rm = TRUE)
3060     comparison_table[comparison_table$Variable == V &
3061         comparison_table$Dataset == "Complete", "Minimum"] <-
3062         min(tapply(datacomplete[, V], datacomplete$cowcode, mean,
3063             na.rm = TRUE), na.rm = TRUE)
3064     comparison_table[comparison_table$Variable == V &
3065         comparison_table$Dataset == "Full", "Maximum"] <-
3066         max(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,

```

```

3067         na.rm = TRUE), na.rm = TRUE)
3068     comparison_table[comparison_table$Variable == V &
3069         comparison_table$Dataset == "Complete", "Maximum"] <-
3070         max(tapply(datacomplete[, V], datacomplete$cowcode, mean,
3071             na.rm = TRUE), na.rm = TRUE)
3072     comparison_table[comparison_table$Variable == V &
3073         comparison_table$Dataset == "Full", "Skew"] <-
3074         Skew(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,
3075             na.rm = TRUE)[is.finite(tapply(autocracy_data[, V],
3076                 autocracy_data$cowcode, mean,
3077                     na.rm = TRUE))])
3078     comparison_table[comparison_table$Variable == V &
3079         comparison_table$Dataset == "Complete", "Skew"] <-
3080         Skew(tapply(datacomplete[, V], datacomplete$cowcode, mean, na.rm = TRUE),
3081             na.rm = TRUE)
3082     #     Because some country means for GDP growth are NaNs, the GDP growth
3083     #     would also be NaN if these are not removed. The values are removed
3084     #     using is.finite(), but doing so distorts the estimated kurtosis.
3085     comparison_table[comparison_table$Variable == V &
3086         comparison_table$Dataset == "Full", "Kurt"] <-
3087         Kurt(tapply(autocracy_data[, V], autocracy_data$cowcode, mean,
3088             na.rm = TRUE)[is.finite(tapply(autocracy_data[, V],
3089                 autocracy_data$cowcode, mean,
3090                     na.rm = TRUE))])
3091     comparison_table[comparison_table$Variable == V &
3092         comparison_table$Dataset == "Complete", "Kurt"] <-
3093         Kurt(tapply(datacomplete[, V], datacomplete$cowcode, mean,
3094             na.rm = TRUE), na.rm = TRUE)
3095     full_intercept_model <- lmer(autocracy_data[, V] ~ (1|cowcode),
3096         data = autocracy_data)
3097     full_intercept_model_ICC_frame <-
3098         as.data.frame(VarCorr(full_intercept_model))
3099     comparison_table[comparison_table$Variable == V &
3100         comparison_table$Dataset == "Full", "ICC"] <-
3101         full_intercept_model_ICC_frame[1, 4] /
3102         (full_intercept_model_ICC_frame[1, 4] +
3103             full_intercept_model_ICC_frame[2, 4])
3104     complete_intercept_model <- lmer(datacomplete[, V] ~ (1|cowcode),
3105         data = datacomplete)
3106     complete_intercept_model_ICC_frame <-
3107         as.data.frame(VarCorr(complete_intercept_model))
3108     comparison_table[comparison_table$Variable == V &
3109         comparison_table$Dataset == "Complete", "ICC"] <-
3110         complete_intercept_model_ICC_frame[1, 4] /
3111         (complete_intercept_model_ICC_frame[1, 4] +
3112             complete_intercept_model_ICC_frame[2, 4])
3113 }
3114
3115 ##### Bivariate analyses #####
3116
3117 # Collect over-time country means
3118 countrymean_table <-
3119     tibble(cowcode = latent_personalism_complete_countrymeans$countrycode,
3120         latent_personalism = latent_personalism_complete_countrymeans$
3121             countrymean,
3122         freedom_assembly = free_assemb_complete_countrymeans$countrymean,
3123         freedom_move = free_move_complete_countrymeans$countrymean,
3124         protection_life_phys = complete_life_phys_countrymeans$countrymean,
3125         freedom_expression = free_expr_complete_countrymeans$countrymean,

```

```

3126     log_population = log_population_complete_countrymeans$countrymean,
3127     log_GDP_per_cap = GDP_cap_complete_countrymeans$countrymean,
3128     GDP_growth = GDP_growth_complete_countrymeans$countrymean,
3129     rigour_impartiality = rigour_impartiality_complete_countrymeans$
3130     countrymean,
3131     political_violence = political_violence_complete_countrymeans$
3132     countrymean,
3133     personal_regime = tapply(datacomplete$gwf_personal,
3134     datacomplete$cowcode, mean),
3135     party_regime = tapply(datacomplete$gwf_party,
3136     datacomplete$cowcode, mean),
3137     military_regime = tapply(datacomplete$gwf_military,
3138     datacomplete$cowcode, mean),
3139     monarch_regime = tapply(datacomplete$gwf_monarch,
3140     datacomplete$cowcode, mean),
3141     international_conflict = tapply(datacomplete$e_miinteco,
3142     datacomplete$cowcode, mean),
3143     internal_conflict = tapply(datacomplete$e_miinterc,
3144     datacomplete$cowcode, mean) )
3145
3146 # Calculate correlations between over-time country means
3147 cor(countrymean_table[, 2:17])
3148
3149 # establish significance threshold
3150 a <- exp((2*1.96)/sqrt(109-3))
3151 (a - 1) / (a + 1)
3152
3153 # Calculate intra-class correlations with year as class
3154
3155 # Personalism
3156 personalism_ICC_model <- lmer(latent_personalism ~ (1|cowcode),
3157     data = datacomplete)
3158 personalism_ICC_frame <-
3159     as.data.frame(VarCorr(personalism_ICC_model))
3160 personalism_ICC_frame[1, 4] / (personalism_ICC_frame[1, 4] +
3161     personalism_ICC_frame[2, 4])
3162
3163 # Freedom of movement
3164 free_move_ICC_model <- lmer(freedom_movement ~ (1|cowcode),
3165     data = datacomplete)
3166 free_move_ICC_frame <-
3167     as.data.frame(VarCorr(free_move_ICC_model))
3168 free_move_ICC_frame[1, 4] / (free_move_ICC_frame[1, 4] +
3169     free_move_ICC_frame[2, 4])
3170
3171 # Freedom of assembly
3172 free_assemb_ICC_model <- lmer(v2caassemb ~ (1|cowcode), data = datacomplete)
3173 free_assemb_ICC_frame <- as.data.frame(VarCorr(free_assemb_ICC_model))
3174 free_assemb_ICC_frame[1, 4] / (free_assemb_ICC_frame[1, 4] +
3175     free_assemb_ICC_frame[2, 4])
3176
3177 # Freedom of expression
3178 free_expr_ICC_model <- lmer(free_expr_x100 ~ (1|cowcode), data = datacomplete)
3179 free_expr_ICC_frame <- as.data.frame(VarCorr(free_expr_ICC_model))
3180 free_expr_ICC_frame[1, 4] / (free_expr_ICC_frame[1, 4] +
3181     free_expr_ICC_frame[2, 4])
3182
3183 # Protection of life and physical integrity
3184 life_phys_ICC_model <- lmer(life_phys_x100 ~ (1|cowcode), data = datacomplete)

```



```

3185 life_phys_ICC_frame <- as.data.frame(VarCorr(life_phys_ICC_model))
3186 life_phys_ICC_frame[1, 4] / (life_phys_ICC_frame[1, 4] +
3187                             life_phys_ICC_frame[2, 4])
3188
3189 # Political violence
3190 political_violence_ICC_model <- lmer(v2caviol ~ (1|cowcode),
3191                                     data = datacomplete)
3192 political_violence_ICC_frame <-
3193   as.data.frame(VarCorr(political_violence_ICC_model))
3194 political_violence_ICC_frame[1, 4] / (political_violence_ICC_frame[1, 4] +
3195                                     political_violence_ICC_frame[2, 4])
3196
3197 # GDP per capita
3198 GDP_cap_ICC_model <- lmer(e_migdppcln ~ (1|cowcode), data = datacomplete)
3199 GDP_cap_ICC_frame <- as.data.frame(VarCorr(GDP_cap_ICC_model))
3200 GDP_cap_ICC_frame[1, 4] / (GDP_cap_ICC_frame[1, 4] + GDP_cap_ICC_frame[2, 4])
3201
3202 # GDP growth
3203 GDP_growth_ICC_model <- lmer(e_migdpgro ~ (1|cowcode), data = datacomplete)
3204 GDP_growth_ICC_frame <- as.data.frame(VarCorr(GDP_growth_ICC_model))
3205 GDP_growth_ICC_frame[1, 4] / (GDP_growth_ICC_frame[1, 4] +
3206                               GDP_growth_ICC_frame[2, 4])
3207
3208 # Rigour and impartiality of the public administration
3209 rig_impact_ICC_model <- lmer(v2clrspct ~ (1|cowcode), data = datacomplete)
3210 rig_impact_ICC_frame <- as.data.frame(VarCorr(rig_impact_ICC_model))
3211 rig_impact_ICC_frame[1, 4] / (rig_impact_ICC_frame[1, 4] +
3212                               rig_impact_ICC_frame[2, 4])
3213
3214 # Population (log 10)
3215 population_ICC_model <- lmer(log10pop ~ (1|cowcode), data = datacomplete)
3216 population_ICC_frame <- as.data.frame(VarCorr(population_ICC_model))
3217 population_ICC_frame[1, 4] / (population_ICC_frame[1, 4] +
3218                               population_ICC_frame[2, 4])
3219
3220 #####
3221 ##### REGRESSION MODELLING #####
3222 #####
3223
3224 ##### Freedom of expression #####
3225
3226 # Intercept-only model
3227 free_expr_intercept <- lmer(free_expr_x100 ~ (1|cowcode), data = datacomplete)
3228
3229 # Calculate clustered standard errors, t-values, p-values and confidence
3230 # intervals
3231 coef_test(free_expr_intercept, vcov = "CR2", test = "Satterthwaite",
3232           coefs = "All", p_values = TRUE)
3233 conf_int(free_expr_intercept, vcov = "CR2", level = 0.9875,
3234          test = "Satterthwaite", coefs = "All", p_values = TRUE)
3235
3236 # Retrieve variance decomposition
3237 free_expr_intercept_ICC_frame <- as.data.frame(VarCorr(free_expr_intercept))
3238 free_expr_intercept_ICC_frame
3239
3240 # Calculate deviance
3241 deviance_free_expr_intercept <- deviance(lmer(free_expr_x100 ~ (1|cowcode),
3242                                               data = datacomplete,
3243                                               REML = FALSE))

```

```

3244
3245 #      ICC
3246 free_expr_intercept_ICC_frame[1, 4] / (free_expr_intercept_ICC_frame[1, 4] +
3247                                       free_expr_intercept_ICC_frame[2, 4])
3248
3249 # Model with year-fixed effects
3250 free_expr_yearfixed <- lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year),
3251                             data = datacomplete)
3252
3253 # Calculate clustered standard errors, t-values, p-values and confidence
3254 # intervals
3255 coef_test(free_expr_yearfixed, vcov = "CR2", test = "Satterthwaite",
3256           coefs = "All", p_values = TRUE)
3257 conf_int(free_expr_yearfixed, vcov = "CR2", level = 0.9875,
3258          test = "Satterthwaite", coefs = "All", p_values = TRUE)
3259
3260 # Retrieve variance decomposition
3261 free_expr_yearfixed_ICC_frame <- as.data.frame(VarCorr(free_expr_yearfixed))
3262 free_expr_yearfixed_ICC_frame
3263
3264 # Calculate R-squared compared to intercept-only model
3265 1 - sum(free_expr_yearfixed_ICC_frame[,4])/
3266     sum(free_expr_intercept_ICC_frame[,4])
3267
3268 #      ICC
3269 free_expr_yearfixed_ICC_frame[1, 4] / (free_expr_yearfixed_ICC_frame[1, 4] +
3270                                       free_expr_yearfixed_ICC_frame[2, 4])
3271
3272 #      LR-test
3273
3274 #      LR-test against intercept-only model
3275 deviance_free_expr_yearfixed <- deviance(lmer(free_expr_x100 ~ (1|cowcode) +
3276                                               as.factor(year),
3277                                               data = datacomplete,
3278                                               REML = FALSE))
3279 deviance_free_expr_intercept - deviance_free_expr_yearfixed
3280 pchisq(deviance_free_expr_intercept - deviance_free_expr_yearfixed,
3281        df = 54, lower.tail = FALSE)
3282
3283 # Model with personalism
3284 free_expr_personalism <- lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year) +
3285                               latent_personalism, data = datacomplete)
3286
3287 # Calculate clustered standard errors, t-values, p-values and confidence
3288 # intervals
3289 coef_test(free_expr_personalism, vcov = "CR2", test = "Satterthwaite",
3290           coefs = "All", p_values = TRUE)
3291 conf_int(free_expr_personalism, vcov = "CR2", level = 0.9875,
3292          test = "Satterthwaite", coefs = "All", p_values = TRUE)
3293
3294 # Retrieve variance decomposition
3295 free_expr_personalism_ICC_frame <- as.data.frame(VarCorr(free_expr_personalism))
3296 free_expr_personalism_ICC_frame
3297
3298 # Calculate R-squared compared to intercept-only model
3299 1 - sum(free_expr_personalism_ICC_frame[,4])/
3300     sum(free_expr_intercept_ICC_frame[,4])
3301
3302 #      ICC

```

```

3303 free_expr_personalism_ICC_frame[1, 4] /
3304   (free_expr_personalism_ICC_frame[1, 4] +
3305     free_expr_personalism_ICC_frame[2, 4])
3306
3307 # Removing year-fixed effects
3308 free_expr_personalism_no_yearfixed <- lmer(free_expr_x100 ~ (1|cowcode) +
3309                                           latent_personalism,
3310                                           data = datacomplete)
3311 coef_test(free_expr_personalism_no_yearfixed, vcov = "CR2",
3312           test = "Satterthwaite", coefs = "All", p_values = TRUE)
3313
3314 # LR-tests
3315
3316 # LR-test against model with year-fixed effects
3317 deviance_free_expr_personalism <- deviance(lmer(free_expr_x100 ~ (1|cowcode) +
3318                                                as.factor(year) +
3319                                                latent_personalism,
3320                                                data = datacomplete,
3321                                                REML = FALSE))
3322
3323 deviance_free_expr_yearfixed - deviance_free_expr_personalism
3324 pchisq(deviance_free_expr_yearfixed - deviance_free_expr_personalism,
3325        df = 1, lower.tail = FALSE)
3326
3327 # LR-test against model without year-fixed effects
3328 deviance_free_expr_personalism_no_yearfixed <-
3329   deviance(lmer(free_expr_x100 ~ (1|cowcode) + latent_personalism,
3330               data = datacomplete, REML = FALSE))
3331 deviance_free_expr_personalism_no_yearfixed - deviance_free_expr_personalism
3332 pchisq(deviance_free_expr_personalism_no_yearfixed -
3333        deviance_free_expr_personalism, df = 54, lower.tail = FALSE)
3334
3335 # Model with personalism and its square
3336 free_expr_personalism_squared <- lmer(free_expr_x100 ~ (1|cowcode) +
3337                                       as.factor(year) +
3338                                       poly(latent_personalism, 2,
3339                                       raw = TRUE), data = datacomplete)
3340
3341 # Calculate clustered standard errors, t-values, p-values and confidence
3342 # intervals
3343 coef_test(free_expr_personalism_squared, vcov = "CR2", test = "Satterthwaite",
3344           coefs = "All", p_values = TRUE)
3345 conf_int(free_expr_personalism_squared, vcov = "CR2", level = 0.9875,
3346          test = "Satterthwaite", coefs = "All", p_values = TRUE)
3347
3348 # Retrieve variance decomposition
3349 free_expr_personalism_squared_ICC_frame <-
3350   as.data.frame(VarCorr(free_expr_personalism_squared))
3351 free_expr_personalism_squared_ICC_frame
3352
3353 # Calculate R-squared compared to intercept-only model
3354 1 - sum(free_expr_personalism_squared_ICC_frame[,4])/
3355   sum(free_expr_intercept_ICC_frame[,4])
3356
3357 # ICC
3358 free_expr_personalism_squared_ICC_frame[1, 4] /
3359   (free_expr_personalism_squared_ICC_frame[1, 4] +
3360     free_expr_personalism_squared_ICC_frame[2, 4])
3361

```

```

3362 # LR-test
3363
3364 # LR-test against personalism-only model
3365 deviance_free_expr_personalism_squared <-
3366   deviance(lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year) +
3367     poly(latent_personalism, 2, raw = TRUE), data = datacomplete,
3368     REML = FALSE))
3369 deviance_free_expr_personalism - deviance_free_expr_personalism_squared
3370 pchisq(deviance_free_expr_personalism - deviance_free_expr_personalism_squared,
3371   df = 1, lower.tail = FALSE)
3372
3373 # LR-test against model without year-fixed effects
3374 deviance_free_expr_personalism_squared_no_yearfixed <-
3375   deviance(lmer(free_expr_x100 ~ (1|cowcode) + poly(latent_personalism, 2,
3376     raw = TRUE),
3377     data = datacomplete, REML = FALSE))
3378 deviance_free_expr_personalism_squared_no_yearfixed -
3379   deviance_free_expr_personalism_squared
3380 pchisq(deviance_free_expr_personalism_squared_no_yearfixed -
3381   deviance_free_expr_personalism_squared, df = 54, lower.tail = FALSE)
3382
3383 # Compare level 1 residuals with linear and quadratic curves
3384
3385 # Quadratic model
3386 ggplot(datacomplete, aes(x = latent_personalism,
3387   y = resid(free_expr_personalism_squared))) +
3388   geom_point(colour = "darkblue") +
3389   geom_smooth(method = lm, formula = y ~ x, colour = "black",
3390     linetype = "dashed", se = FALSE) + blue_light +
3391   labs(x = "Personalism", y = "Level 1 residual")
3392
3393 # Linear model
3394 ggplot(datacomplete, aes(x = latent_personalism,
3395   y = resid(free_expr_personalism))) +
3396   geom_point(colour = "darkblue") +
3397   geom_smooth(method = lm, formula = y ~ x, colour = "black",
3398     linetype = "dashed", se = FALSE) + blue_light +
3399   labs(x = "Personalism", y = "Level 1 residual")
3400
3401 # Model with personalism and controls
3402 free_expr_control <- lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year) +
3403   latent_personalism + lag_e_miinteco +
3404   lag_e_miinterc + lag_e_migdppc1n + lag_e_migdppgro +
3405   lag_v2caviol + lag_log10pop + gwf_monarch +
3406   gwf_military + gwf_party, data = datacomplete)
3407
3408 # Calculate clustered standard errors, t-values, p-values and confidence
3409 # intervals
3410 coef_test(free_expr_control, vcov = "CR2", test = "Satterthwaite",
3411   coefs = "All", p_values = TRUE)
3412 conf_int(free_expr_control, vcov = "CR2", level = 0.9875,
3413   test = "Satterthwaite", coefs = "All", p_values = TRUE)
3414
3415 # Retrieve variance decomposition
3416 free_expr_control_ICC_frame <- as.data.frame(VarCorr(free_expr_control))
3417 free_expr_control_ICC_frame
3418
3419 # Calculate R-squared compared to intercept-only model
3420 1 - sum(free_expr_control_ICC_frame[,4])/ sum(free_expr_intercept_ICC_frame[,4])

```

```

3421
3422 #      ICC
3423 free_expr_control_ICC_frame[1, 4] / (free_expr_control_ICC_frame[1, 4] +
3424                                     free_expr_control_ICC_frame[2, 4])
3425
3426 #      Removing year-fixed effects
3427 free_expr_control_no_yearfixed <- lmer(free_expr_x100 ~ (1|cowcode) +
3428                                     latent_personalism +
3429                                     lag_e_miinteco + lag_e_miinterc +
3430                                     lag_e_migdppln + lag_e_migdpgro +
3431                                     lag_v2caviol + lag_log10pop +
3432                                     gwf_monarch + gwf_military + gwf_party,
3433                                     data = datacomplete)
3434 coef_test(free_expr_control_no_yearfixed, vcov = "CR2", test = "Satterthwaite",
3435           coefs = "All", p_values = TRUE)
3436
3437 #      LR-test
3438
3439 #      LR-test against personalism-only model
3440 deviance_free_expr_control <-
3441   deviance(lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year) +
3442               latent_personalism + lag_e_miinteco + lag_e_miinterc +
3443               lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
3444               lag_log10pop + gwf_monarch + gwf_military + gwf_party,
3445               data = datacomplete, REML = FALSE))
3446 deviance_free_expr_personalism - deviance_free_expr_control
3447 pchisq(deviance_free_expr_personalism - deviance_free_expr_control,
3448        df = 9, lower.tail = FALSE)
3449
3450 #      LR-test against model without year-fixed effects
3451 deviance_free_expr_control_no_yearfixed <-
3452   deviance(lmer(free_expr_x100 ~ (1|cowcode) + latent_personalism +
3453               lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
3454               lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
3455               gwf_military + gwf_party, data = datacomplete, REML = FALSE))
3456 deviance_free_expr_control_no_yearfixed - deviance_free_expr_control
3457 pchisq(deviance_free_expr_control_no_yearfixed - deviance_free_expr_control,
3458        df = 54, lower.tail = FALSE)
3459
3460 #      Residual diagnostics
3461
3462 #      Add predicted values to dataset
3463 datacomplete$free_expr_predict <- predict(free_expr_control)
3464
3465 #      Add residuals to dataset
3466
3467 #      Calculate level 1 residuals
3468 datacomplete$free_expr_resid_level_1 <- residuals(free_expr_control)
3469
3470 #      Standardise level 1 residuals
3471 for(c in unique(datacomplete$cowcode)) {
3472   datacomplete$free_expr_country_resid_sds[datacomplete$cowcode == c] <-
3473     sd(datacomplete$free_expr_resid_level_1[datacomplete$cowcode == c])
3474 }
3475 datacomplete$free_expr_stand_resid_level_1 <-
3476   datacomplete$free_expr_resid_level_1 /
3477   datacomplete$free_expr_country_resid_sds
3478
3479 #      Calculate level 2 residuals

```

```

3480 free_expr_resid_level_2 <- as.data.frame(ranef(free_expr_control))
3481 free_expr_resid_level_2$cowcode <-
3482   as.numeric(levels(free_expr_resid_level_2$grp))[free_expr_resid_level_2$grp]
3483
3484 # Standardise level 2 residuals
3485 free_expr_resid_level_2$free_expr_stand_resid_level_2 <-
3486   free_expr_resid_level_2$condval /
3487   free_expr_control_ICC_frame$sdcor[1]
3488
3489 # Add level 2 residuals to dataset
3490 datacomplete <- left_join(datacomplete,
3491   free_expr_resid_level_2[, c("cowcode", "condval",
3492     "free_expr_stand_resid_level_2")],
3493   by = c("cowcode"))
3494 datacomplete <- rename(datacomplete, free_expr_resid_level_2 = condval)
3495
3496 # Autocorrelation
3497
3498 # Create dataframe with just the residuals
3499 free_expr_resid_pdata <-
3500   pdata.frame(datacomplete[, c("year", "cowcode", "free_expr_resid_level_1")],
3501     index = c("cowcode", "year"))
3502 # Create lagged residuals
3503 for (j in 1:20) {
3504   # Create the name of the lagged variable
3505   var <- paste("free_expr_lag", j, "_resid_level_1", sep = "")
3506   # Generate the lagged variable via plm's implementation of "lag"
3507   free_expr_resid_pdata[, var] <-
3508     plm::lag(free_expr_resid_pdata$free_expr_resid_level_1, k= j)
3509 }
3510
3511 # Calculate correlations between present values and lags
3512 autocors_free_expr_resid_level_1 <-
3513   cor(free_expr_resid_pdata$free_expr_resid_level_1,
3514     free_expr_resid_pdata[, 4:ncol(free_expr_resid_pdata)],
3515     use = "complete.obs")
3516
3517 # Reshape the correlations into a more workable format
3518 autocors_free_expr_resid_level_1 <-
3519   tibble(Lag = 1:20, Correlation = t(autocors_free_expr_resid_level_1))
3520
3521 # Create an autocorrelation plot
3522 ggplot(data = autocors_free_expr_resid_level_1, aes(x = Lag, y = Correlation)) +
3523   geom_col(fill = "lightblue") + blue_light
3524
3525 # Linearity
3526
3527 # Plot predicted values against level 1 residuals
3528 ggplot(datacomplete, aes(x = free_expr_predict,
3529   y = free_expr_stand_resid_level_1)) +
3530   geom_point(colour = "lightblue") +
3531   geom_smooth(method = lm, formula = y ~ x, colour = "black",
3532     linetype = "dashed", se = FALSE) + blue_light +
3533   labs(x = "Predicted value", y = "Level 1 residual (standardised)")
3534
3535 # Plot predicted values from fixed effects against level 2 residuals
3536 ggplot(datacomplete, aes(x = free_expr_predict - free_expr_resid_level_2,
3537   y = free_expr_stand_resid_level_2)) +
3538   geom_point(colour = "lightblue") +

```

```

3539 geom_smooth(method = lm, formula = y ~ x, colour = "black",
3540             linetype = "dashed", se = FALSE) + blue_light +
3541 labs(x = "Fixed part predictions", y = "Level 2 residual (standardised)")
3542
3543 # Normality
3544
3545 # Assign residual levels to the residuals for combination
3546 rbind_resids_level_2 <-
3547   rename(free_expr_resid_level_2[, c("cowcode",
3548                                     "free_expr_stand_resid_level_2")],
3549          resid = free_expr_stand_resid_level_2)
3550 rbind_resids_level_2$level <- "Level 2"
3551 rbind_resids_level_1 <-
3552   rename(datacomplete[, c("cowcode", "free_expr_stand_resid_level_1")],
3553          resid = free_expr_stand_resid_level_1)
3554 rbind_resids_level_1$level <- "Level 1"
3555
3556 # Create long dataset of all residuals
3557 free_expr_resids <- rbind(rbind_resids_level_1, rbind_resids_level_2)
3558 rm(list = c("rbind_resids_level_1", "rbind_resids_level_2"))
3559
3560 # Plot residuals in dual QQ-plot
3561 withr::with_options(
3562   list(ggplot2.discrete.colour = c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
3563                                     "#0072B2", "#D55E00", "#CC79A7")),
3564   print(ggplot(data = free_expr_resids, aes(sample = resid,
3565                                             colour = as.factor(level))) +
3566         geom_qq() + geom_qq_line() + blue_light +
3567         scale_colour_discrete(name = "Residual level") +
3568         labs(y = "Standardised residuals",
3569              x = "Reference normal distribution")) )
3570
3571 # Plot residuals in single QQ-plots
3572
3573 # Level 1
3574 ggplot(data = free_expr_resids[free_expr_resids$level == "Level 1", ],
3575        aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
3576   blue_light + labs(y = "Standardised residuals",
3577                    x = "Reference normal distribution")
3578
3579 # Level 2
3580 ggplot(data = free_expr_resids[free_expr_resids$level == "Level 2", ],
3581        aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
3582   blue_light + labs(y = "Standardised residuals", x = "Reference normal
3583   distribution")
3584
3585 # Model with rigour and impartiality of the public administration
3586 free_expr_rig_impact <- lmer(free_expr_x100 ~ (1|cowcode) + as.factor(year) +
3587                             latent_personalism + v2clrspct + lag_e_miinteco +
3588                             lag_e_miinterc + lag_e_migdppcln +
3589                             lag_e_migdpgro + lag_v2caviol + lag_log10pop +
3590                             gwf_monarch + gwf_military + gwf_party,
3591                             data = datacomplete)
3592
3593 # Calculate clustered standard errors, t-values, p-values and confidence
3594 # intervals
3595 coef_test(free_expr_rig_impact, vcov = "CR2", test = "Satterthwaite",
3596           coefs = "All", p_values = TRUE)
3597 conf_int(free_expr_rig_impact, vcov = "CR2", level = 0.9875,

```

```

3598         test = "Satterthwaite", coefs = "All", p_values = TRUE)
3599
3600 # Retrieve variance decomposition
3601 free_expr_rig_impact_ICC_frame <- as.data.frame(VarCorr(free_expr_rig_impact))
3602 free_expr_rig_impact_ICC_frame
3603
3604 # Removing year-fixed effects
3605 free_expr_rig_impact_no_yearfixed <- lmer(free_expr_x100 ~ (1|cowcode) +
3606                                           latent_personalism + v2clrspct +
3607                                           lag_e_miinteco + lag_e_miinterc +
3608                                           lag_e_migdppln + lag_e_migdpgro +
3609                                           lag_v2caviol + lag_log10pop +
3610                                           gwf_monarch + gwf_military +
3611                                           gwf_party, data = datacomplete)
3612 coef_test(free_expr_rig_impact_no_yearfixed, vcov = "CR2",
3613           test = "Satterthwaite", coefs = "All", p_values = TRUE)
3614
3615 # LR-test and ICC
3616
3617 # LR-test against personalism-only model
3618 deviance_free_expr_rig_impact <- deviance(lmer(free_expr_x100 ~ (1|cowcode) +
3619                                               as.factor(year) +
3620                                               latent_personalism +
3621                                               v2clrspct + lag_e_miinteco +
3622                                               lag_e_miinterc +
3623                                               lag_e_migdppln +
3624                                               lag_e_migdpgro +
3625                                               lag_v2caviol + lag_log10pop +
3626                                               gwf_monarch + gwf_military +
3627                                               gwf_party,
3628                                               data = datacomplete, REML = FALSE))
3629 deviance_free_expr_control - deviance_free_expr_rig_impact
3630 pchisq(deviance_free_expr_control - deviance_free_expr_rig_impact,
3631        df = 9, lower.tail = FALSE)
3632
3633 # LR-test against model without year-fixed effects
3634 deviance_free_expr_rig_impact_no_yearfixed <-
3635   deviance(lmer(free_expr_x100 ~ (1|cowcode) + latent_personalism + v2clrspct +
3636               lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
3637               lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
3638               gwf_military + gwf_party, data = datacomplete, REML = FALSE))
3639 deviance_free_expr_rig_impact_no_yearfixed - deviance_free_expr_rig_impact
3640 pchisq(deviance_free_expr_rig_impact_no_yearfixed -
3641        deviance_free_expr_rig_impact, df = 54, lower.tail = FALSE)
3642
3643 # ICC
3644 free_expr_rig_impact_ICC_frame[1, 4] / (free_expr_rig_impact_ICC_frame[1, 4] +
3645                                         free_expr_rig_impact_ICC_frame[2, 4])
3646
3647 ##### Freedom of assembly #####
3648
3649 # Intercept-only model
3650 free_assemb_intercept <- lmer(v2caassemb ~ (1|cowcode), data = datacomplete)
3651
3652 # Calculate clustered standard errors, t-values, p-values and confidence
3653 # intervals
3654 coef_test(free_assemb_intercept, vcov = "CR2", test = "Satterthwaite",
3655           coefs = "All", p_values = TRUE)
3656 conf_int(free_assemb_intercept, vcov = "CR2", level = 0.9875,

```



```

3657         test = "Satterthwaite", coefs = "All", p_values = TRUE)
3658
3659 # Retrieve variance decomposition
3660 free_assemb_intercept_ICC_frame <- as.data.frame(VarCorr(free_assemb_intercept))
3661 free_assemb_intercept_ICC_frame
3662
3663 # Calculate deviance
3664 deviance_free_assemb_intercept <- deviance(lmer(v2caassemb ~ (1|cowcode),
3665                                             data = datacomplete,
3666                                             REML = FALSE))
3667
3668 # ICC
3669 free_assemb_intercept_ICC_frame[1, 4] /
3670 (free_assemb_intercept_ICC_frame[1, 4] +
3671  free_assemb_intercept_ICC_frame[2, 4])
3672
3673 # Model with year-fixed effects
3674 free_assemb_yearfixed <- lmer(v2caassemb ~ (1|cowcode) + as.factor(year),
3675                             data = datacomplete)
3676
3677 # Calculate clustered standard errors, t-values, p-values and confidence
3678 # intervals
3679 coef_test(free_assemb_yearfixed, vcov = "CR2", test = "Satterthwaite",
3680          coefs = "All", p_values = TRUE)
3681 conf_int(free_assemb_yearfixed, vcov = "CR2", level = 0.9875,
3682         test = "Satterthwaite", coefs = "All", p_values = TRUE)
3683
3684 # Retrieve variance decomposition
3685 free_assemb_yearfixed_ICC_frame <- as.data.frame(VarCorr(free_assemb_yearfixed))
3686 free_assemb_yearfixed_ICC_frame
3687
3688 # Calculate R-squared compared to intercept-only model
3689 1 - sum(free_assemb_yearfixed_ICC_frame[,4])/
3690 sum(free_assemb_intercept_ICC_frame[,4])
3691
3692 # ICC
3693 free_assemb_yearfixed_ICC_frame[1, 4] /
3694 (free_assemb_yearfixed_ICC_frame[1, 4] +
3695  free_assemb_yearfixed_ICC_frame[2, 4])
3696
3697 # LR-test
3698
3699 # LR-test against intercept-only model
3700 deviance_free_assemb_yearfixed <- deviance(lmer(v2caassemb ~ (1|cowcode) +
3701                                             as.factor(year),
3702                                             data = datacomplete,
3703                                             REML = FALSE))
3704 deviance_free_assemb_intercept - deviance_free_assemb_yearfixed
3705 pchisq(deviance_free_assemb_intercept - deviance_free_assemb_yearfixed,
3706        df = 54, lower.tail = FALSE)
3707
3708 # Model with personalism
3709 free_assemb_personalism <- lmer(v2caassemb ~ (1|cowcode) + as.factor(year) +
3710                             latent_personalism, data = datacomplete)
3711
3712 # Calculate clustered standard errors, t-values, p-values and confidence
3713 # intervals
3714 coef_test(free_assemb_personalism, vcov = "CR2", test = "Satterthwaite",
3715          coefs = "All", p_values = TRUE)

```

```

3716 conf_int(free_assemb_personalism, vcov = "CR2", level = 0.9875,
3717           test = "Satterthwaite", coefs = "All", p_values = TRUE)
3718
3719 # Retrieve variance decomposition
3720 free_assemb_personalism_ICC_frame <-
3721   as.data.frame(VarCorr(free_assemb_personalism))
3722 free_assemb_personalism_ICC_frame
3723
3724 # Calculate R-squared compared to intercept-only model
3725 1 - sum(free_assemb_personalism_ICC_frame[,4])/
3726   sum(free_assemb_intercept_ICC_frame[,4])
3727
3728 # ICC
3729 free_assemb_personalism_ICC_frame[1, 4] /
3730   (free_assemb_personalism_ICC_frame[1, 4] +
3731     free_assemb_personalism_ICC_frame[2, 4])
3732
3733 # Removing year-fixed effects
3734 free_assemb_rig_personalism_no_yearfixed <- lmer(v2caassemb ~ (1|cowcode) +
3735                                               latent_personalism,
3736                                               data = datacomplete)
3737 coef_test(free_assemb_rig_personalism_no_yearfixed, vcov = "CR2",
3738           test = "Satterthwaite", coefs = "All", p_values = TRUE)
3739
3740 # LR-test
3741
3742 # LR-test against model with year-fixed effects
3743 deviance_free_assemb_personalism <- deviance(lmer(v2caassemb ~ (1|cowcode) +
3744                                               as.factor(year) +
3745                                               latent_personalism,
3746                                               data = datacomplete,
3747                                               REML = FALSE))
3748
3749 deviance_free_assemb_yearfixed - deviance_free_assemb_personalism
3750 pchisq(deviance_free_assemb_yearfixed - deviance_free_assemb_personalism,
3751        df = 1, lower.tail = FALSE)
3752
3753 # LR-test against model without year-fixed effects
3754 deviance_free_assemb_personalism_no_yearfixed <-
3755   deviance(lmer(v2caassemb ~ (1|cowcode) + latent_personalism,
3756               data = datacomplete, REML = FALSE))
3757 deviance_free_assemb_personalism_no_yearfixed - deviance_free_assemb_personalism
3758 pchisq(deviance_free_assemb_personalism_no_yearfixed -
3759        deviance_free_assemb_personalism,
3760        df = 54, lower.tail = FALSE)
3761
3762 # Model with personalism and controls
3763 free_assemb_control <- lmer(v2caassemb ~ (1|cowcode) + as.factor(year) +
3764                             latent_personalism + lag_e_miinteco +
3765                             lag_e_miinterc + lag_e_migdppcln +
3766                             lag_e_migdpgro + lag_v2caviol + lag_log10pop +
3767                             gwf_monarch + gwf_military + gwf_party,
3768                             data = datacomplete)
3769
3770 # Calculate clustered standard errors, t-values, p-values and confidence
3771 # intervals
3772 coef_test(free_assemb_control, vcov = "CR2", test = "Satterthwaite",
3773           coefs = "All", p_values = TRUE)
3774 conf_int(free_assemb_control, vcov = "CR2", level = 0.9875,

```

```

3775     test = "Satterthwaite", coefs = "All", p_values = TRUE)
3776
3777 # Retrieve variance decomposition
3778 free_assemb_control_ICC_frame <- as.data.frame(VarCorr(free_assemb_control))
3779 free_assemb_control_ICC_frame
3780
3781 # Calculate R-squared compared to intercept-only model
3782 1 - sum(free_assemb_control_ICC_frame[,4])/
3783     sum(free_assemb_intercept_ICC_frame[,4])
3784
3785 # ICC
3786 free_assemb_control_ICC_frame[1, 4] / (free_assemb_control_ICC_frame[1, 4] +
3787     free_assemb_control_ICC_frame[2, 4])
3788
3789 # Removing year-fixed effects
3790 free_assemb_control_no_yearfixed <-
3791     lmer(v2caassemb ~ (1|cowcode) + latent_personalism + lag_e_miinteco +
3792         lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
3793         lag_log10pop + gwf_monarch + gwf_military + gwf_party,
3794         data = datacomplete)
3795 coef_test(free_assemb_control_no_yearfixed, vcov = "CR2",
3796     test = "Satterthwaite", coefs = "All", p_values = TRUE)
3797
3798 # LR-test
3799
3800 # LR-test against personalism-only model
3801 deviance_free_assemb_control <-
3802     deviance(lmer(v2caassemb ~ (1|cowcode) + as.factor(year) +
3803         latent_personalism + lag_e_miinteco + lag_e_miinterc +
3804         lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
3805         lag_log10pop + gwf_monarch + gwf_military + gwf_party,
3806         data = datacomplete, REML = FALSE))
3807 deviance_free_assemb_personalism - deviance_free_assemb_control
3808 pchisq(deviance_free_assemb_personalism - deviance_free_assemb_control,
3809     df = 9, lower.tail = FALSE)
3810
3811 # LR-test against model without year-fixed effects
3812 deviance_free_assemb_control_no_yearfixed <-
3813     deviance(lmer(v2caassemb ~ (1|cowcode) + latent_personalism +
3814         lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
3815         lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
3816         gwf_military + gwf_party, data = datacomplete, REML = FALSE))
3817 deviance_free_assemb_control_no_yearfixed - deviance_free_assemb_control
3818 pchisq(deviance_free_assemb_control_no_yearfixed - deviance_free_assemb_control,
3819     df = 54, lower.tail = FALSE)
3820
3821 # Residual diagnostics
3822
3823 # Add predicted values to dataset
3824 datacomplete$free_assemb_predict <- predict(free_assemb_control)
3825
3826 # Add residuals to dataset
3827
3828 # Calculate level 1 residuals
3829 datacomplete$free_assemb_resid_level_1 <- residuals(free_assemb_control)
3830
3831 # Standardise level 1 residuals
3832 for(c in unique(datacomplete$cowcode)) {
3833     datacomplete$free_assemb_country_resid_sds[datacomplete$cowcode == c] <-

```

```

3834     sd(datacomplete$free_assemb_resid_level_1[datacomplete$cowcode == c])
3835 }
3836 datacomplete$free_assemb_stand_resid_level_1 <-
3837   datacomplete$free_assemb_resid_level_1 /
3838   datacomplete$free_assemb_country_resid_sds
3839
3840 #       Calculate level 2 residuals
3841 free_assemb_resid_level_2 <- as.data.frame(ranef(free_assemb_control))
3842 free_assemb_resid_level_2$cowcode <-
3843   as.numeric(levels(free_assemb_resid_level_2$grp))[free_assemb_resid_level_2$
3844     grp]
3845
3846 #       Standardise level 2 residuals
3847 free_assemb_resid_level_2$free_assemb_stand_resid_level_2 <-
3848   free_assemb_resid_level_2$condval / free_assemb_control_ICC_frame$sdcor[1]
3849
3850 #       Add level 2 residuals to dataset
3851 datacomplete <-
3852   left_join(datacomplete,
3853     free_assemb_resid_level_2[, c("cowcode", "condval",
3854       "free_assemb_stand_resid_level_2")],
3855     by = c("cowcode"))
3856 datacomplete <- rename(datacomplete, free_assemb_resid_level_2 = condval)
3857
3858 #       Autocorrelation
3859
3860 #       Create dataframe with just the residuals
3861 free_assemb_resid_pdata <-
3862   pdata.frame(datacomplete[, c("year", "cowcode", "free_assemb_resid_level_1")],
3863     index = c("cowcode", "year"))
3864 #       Create lagged residuals
3865 for (j in 1:20) {
3866   # Create the name of the lagged variable
3867   var <- paste("free_assemb_lag", j, "_resid_level_1", sep = "")
3868   # Generate the lagged variable via plm's implementation of "lag"
3869   free_assemb_resid_pdata[, var] <-
3870     plm::lag(free_assemb_resid_pdata$free_assemb_resid_level_1, k= j)
3871 }
3872
3873 #       Calculate correlations between present values and lags
3874 autocors_free_assemb_resid_level_1 <-
3875   cor(free_assemb_resid_pdata$free_assemb_resid_level_1,
3876     free_assemb_resid_pdata[, 4:ncol(free_assemb_resid_pdata)],
3877     use = "complete.obs")
3878
3879 #       Reshape the correlations into a more workable format
3880 autocors_free_assemb_resid_level_1 <-
3881   tibble(Lag = 1:20, Correlation = t(autocors_free_assemb_resid_level_1))
3882
3883 #       Create an autocorrelation plot
3884 ggplot(data = autocors_free_assemb_resid_level_1,
3885   aes(x = Lag, y = Correlation)) + geom_col(fill = "lightblue") +
3886   blue_light
3887
3888 #       Linearity
3889
3890 #       Plot predicted values against level 1 residuals
3891 ggplot(datacomplete, aes(x = free_assemb_predict,
3892   y = free_assemb_stand_resid_level_1)) +

```

```

3893   geom_point(colour = "lightblue") +
3894   geom_smooth(method = lm, formula = y ~ x, colour = "black",
3895             linetype = "dashed", se = FALSE) +
3896   blue_light + labs(x = "Predicted value",
3897                   y = "Level 1 residual (standardised)")
3898
3899   # Plot values predicted by fixed part against level 2 residuals
3900   ggplot(datacomplete, aes(x = free_assemb_predict - free_assemb_resid_level_2,
3901                           y = free_assemb_stand_resid_level_2)) +
3902   geom_point(colour = "lightblue") +
3903   geom_smooth(method = lm, formula = y ~ x, colour = "black",
3904             linetype = "dashed", se = FALSE) +
3905   blue_light + labs(x = "Fixed part predictions",
3906                   y = "Level 2 residual (standardised)")
3907
3908   # Normality
3909
3910   # Assign residual levels to the residuals for combination
3911   rbind_resids_level_2 <-
3912     rename(free_assemb_resid_level_2[, c("cowcode",
3913                                         "free_assemb_stand_resid_level_2")],
3914           resid = free_assemb_stand_resid_level_2)
3915   rbind_resids_level_2$level <- "Level 2"
3916   rbind_resids_level_1 <-
3917     rename(datacomplete[, c("cowcode", "free_assemb_stand_resid_level_1")],
3918           resid = free_assemb_stand_resid_level_1)
3919   rbind_resids_level_1$level <- "Level 1"
3920
3921   # Create long dataset of all residuals
3922   free_assemb_resids <- rbind(rbind_resids_level_1, rbind_resids_level_2)
3923   rm(list = c("rbind_resids_level_1", "rbind_resids_level_2"))
3924
3925   # Plot residuals in dual QQ-plot
3926   withr::with_options(
3927     list(ggplot2.discrete.colour = c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
3928                                     "#0072B2", "#D55E00", "#CC79A7")),
3929     print(ggplot(data = free_assemb_resids, aes(sample = resid,
3930                                               colour = as.factor(level))) +
3931           geom_qq() + geom_qq_line() + blue_light +
3932           scale_colour_discrete(name = "Residual level") +
3933           labs(y = "Standardised residuals",
3934               x = "Reference normal distribution"))
3935   )
3936
3937   # Plot residuals in single QQ-plots
3938
3939   # Level 1
3940   ggplot(data = free_assemb_resids[free_assemb_resids$level == "Level 1", ],
3941         aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
3942   blue_light + labs(y = "Standardised residuals",
3943                   x = "Reference normal distribution")
3944
3945   # Level 2
3946   ggplot(data = free_assemb_resids[free_assemb_resids$level == "Level 2", ],
3947         aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
3948   blue_light + labs(y = "Standardised residuals",
3949                   x = "Reference normal distribution")
3950
3951   # Model with rigour and impartiality of the public administration

```

```

3952 free_assemb_rig_impact <-
3953   lmer(v2caassemb ~ (1|cowcode) + as.factor(year) + latent_personalism +
3954     v2clrspct + lag_e_miinteco + lag_e_miinterc + lag_e_migdppc1n +
3955     lag_e_migdpgr0 + lag_v2caviol + lag_log10pop + gwf_monarch +
3956     gwf_military + gwf_party, data = datacomplete)
3957
3958 #   Calculate clustered standard errors, t-values, p-values and confidence
3959 #   intervals
3960 coef_test(free_assemb_rig_impact, vcov = "CR2", test = "Satterthwaite",
3961   coefs = "All", p_values = TRUE)
3962 conf_int(free_assemb_rig_impact, vcov = "CR2", level = 0.9875,
3963   test = "Satterthwaite", coefs = "All", p_values = TRUE)
3964
3965 #   Retrieve variance decomposition
3966 free_assemb_rig_impact_ICC_frame <-
3967   as.data.frame(VarCorr(free_assemb_rig_impact))
3968 free_assemb_rig_impact_ICC_frame
3969
3970 #   Calculate R-squared compared to intercept-only model
3971 1 - sum(free_assemb_rig_impact_ICC_frame[,4])/
3972   sum(free_assemb_intercept_ICC_frame[,4])
3973
3974 #   ICC
3975 free_assemb_rig_impact_ICC_frame[1, 4] /
3976   (free_assemb_rig_impact_ICC_frame[1, 4] +
3977     free_assemb_rig_impact_ICC_frame[2, 4])
3978
3979 #   Removing year-fixed effects
3980 free_assemb_rig_impact_no_yearfixed <-
3981   lmer(v2caassemb ~ (1|cowcode) + latent_personalism + v2clrspct +
3982     lag_e_miinteco + lag_e_miinterc + lag_e_migdppc1n + lag_e_migdpgr0 +
3983     lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
3984     data = datacomplete)
3985 coef_test(free_assemb_rig_impact_no_yearfixed, vcov = "CR2",
3986   test = "Satterthwaite", coefs = "All", p_values = TRUE)
3987
3988 #   LR-test
3989
3990 #   LR-test against model with controls
3991 deviance_free_assemb_rig_impact <-
3992   deviance(lmer(v2caassemb ~ (1|cowcode) + as.factor(year) +
3993     latent_personalism + v2clrspct + lag_e_miinteco +
3994     lag_e_miinterc + lag_e_migdppc1n + lag_e_migdpgr0 +
3995     lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military +
3996     gwf_party, data = datacomplete, REML = FALSE))
3997 deviance_free_assemb_control - deviance_free_assemb_rig_impact
3998 pchisq(deviance_free_assemb_control - deviance_free_assemb_rig_impact,
3999   df = 9, lower.tail = FALSE)
4000
4001 #   LR-test against model without year-fixed effects
4002 deviance_free_assemb_rig_impact_no_yearfixed <-
4003   deviance(lmer(v2caassemb ~ (1|cowcode) + latent_personalism + v2clrspct +
4004     lag_e_miinteco + lag_e_miinterc + lag_e_migdppc1n +
4005     lag_e_migdpgr0 + lag_v2caviol + lag_log10pop + gwf_monarch +
4006     gwf_military + gwf_party, data = datacomplete, REML = FALSE))
4007 deviance_free_assemb_rig_impact_no_yearfixed - deviance_free_assemb_rig_impact
4008 pchisq(deviance_free_assemb_rig_impact_no_yearfixed -
4009   deviance_free_assemb_rig_impact, df = 54, lower.tail = FALSE)
4010

```

```

4011 ##### Protection of life and physical integrity #####
4012
4013 # Intercept-only model
4014 life_phys_intercept <- lmer(life_phys_x100 ~ (1|cowcode), data = datacomplete)
4015
4016 # Calculate clustered standard errors, t-values, p-values and confidence
4017 # intervals
4018 coef_test(life_phys_intercept, vcov = "CR2", test = "Satterthwaite",
4019           coefs = "All", p_values = TRUE)
4020 conf_int(life_phys_intercept, vcov = "CR2", level = 0.9875,
4021          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4022
4023 # Retrieve variance decomposition
4024 life_phys_intercept_ICC_frame <- as.data.frame(VarCorr(life_phys_intercept))
4025 life_phys_intercept_ICC_frame
4026
4027 # Calculate deviance
4028 deviance_life_phys_intercept <- deviance(lmer(life_phys_x100 ~ (1|cowcode),
4029                                             data = datacomplete,
4030                                             REML = FALSE))
4031
4032 # ICC
4033 life_phys_intercept_ICC_frame[1, 4] / (life_phys_intercept_ICC_frame[1, 4] +
4034                                       life_phys_intercept_ICC_frame[2, 4])
4035
4036 # Model with year-fixed effects
4037 life_phys_yearfixed <- lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year),
4038                             data = datacomplete)
4039
4040 # Calculate clustered standard errors, t-values, p-values and confidence
4041 # intervals
4042 coef_test(life_phys_yearfixed, vcov = "CR2", test = "Satterthwaite",
4043           coefs = "All", p_values = TRUE)
4044 conf_int(life_phys_yearfixed, vcov = "CR2", level = 0.9875,
4045          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4046
4047 # Retrieve variance decomposition and REML criterion
4048 life_phys_yearfixed_ICC_frame <- as.data.frame(VarCorr(life_phys_yearfixed))
4049 life_phys_yearfixed_ICC_frame
4050
4051 # Calculate R-squared compared to intercept-only model
4052 1 - sum(life_phys_yearfixed_ICC_frame[,4])/
4053     sum(life_phys_intercept_ICC_frame[,4])
4054
4055 # ICC
4056 life_phys_yearfixed_ICC_frame[1, 4] / (life_phys_yearfixed_ICC_frame[1, 4] +
4057                                       life_phys_yearfixed_ICC_frame[2, 4])
4058
4059 # LR-test and ICC
4060
4061 # LR-test against intercept-only model
4062 deviance_life_phys_yearfixed <- deviance(lmer(life_phys_x100 ~ (1|cowcode) +
4063                                             as.factor(year),
4064                                             data = datacomplete,
4065                                             REML = FALSE))
4066 deviance_life_phys_intercept - deviance_life_phys_yearfixed
4067 pchisq(deviance_life_phys_intercept - deviance_life_phys_yearfixed,
4068        df = 54, lower.tail = FALSE)
4069

```

```

4070 # Model with personalism
4071 life_phys_personalism <- lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) +
4072                             latent_personalism, data = datacomplete)
4073
4074 # Calculate clustered standard errors, t-values, p-values and confidence
4075 # intervals
4076 coef_test(life_phys_personalism, vcov = "CR2", test = "Satterthwaite",
4077           coefs = "All", p_values = TRUE)
4078 conf_int(life_phys_personalism, vcov = "CR2", level = 0.9875,
4079          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4080
4081 # Retrieve variance decomposition
4082 life_phys_personalism_ICC_frame <- as.data.frame(VarCorr(life_phys_personalism))
4083 life_phys_personalism_ICC_frame
4084
4085 # Calculate R-squared compared to intercept-only model
4086 1 - sum(life_phys_personalism_ICC_frame[,4])/
4087     sum(life_phys_intercept_ICC_frame[,4])
4088
4089 # ICC
4090 life_phys_personalism_ICC_frame[1, 4] /
4091   (life_phys_personalism_ICC_frame[1, 4] +
4092    life_phys_personalism_ICC_frame[2, 4])
4093
4094 # Removing year-fixed effects
4095 life_phys_rig_personalism_no_yearfixed <-
4096   lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism, data = datacomplete)
4097 coef_test(life_phys_rig_personalism_no_yearfixed, vcov = "CR2",
4098           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4099
4100 # LR-test
4101
4102 # LR-test against model with year-fixed effects
4103 deviance_life_phys_personalism <-
4104   deviance(lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) +
4105               latent_personalism, data = datacomplete, REML = FALSE))
4106
4107 deviance_life_phys_yearfixed - deviance_life_phys_personalism
4108 pchisq(deviance_life_phys_yearfixed - deviance_life_phys_personalism,
4109        df = 1, lower.tail = FALSE)
4110
4111 # LR-test against model without year-fixed effects
4112 deviance_life_phys_personalism_no_yearfixed <-
4113   deviance(lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism,
4114               data = datacomplete, REML = FALSE))
4115 deviance_life_phys_personalism_no_yearfixed - deviance_life_phys_personalism
4116 pchisq(deviance_life_phys_personalism_no_yearfixed -
4117        deviance_life_phys_personalism, df = 54, lower.tail = FALSE)
4118
4119 # Model with personalism and controls
4120 life_phys_control <-
4121   lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) + latent_personalism +
4122       lag_e_miinteco + lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4123       lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4124       data = datacomplete)
4125
4126 # Calculate clustered standard errors, t-values, p-values and confidence
4127 # intervals
4128 coef_test(life_phys_control, vcov = "CR2", test = "Satterthwaite",

```



```

4129     coefs = "All", p_values = TRUE)
4130 conf_int(life_phys_control, vcov = "CR2", level = 0.9875,
4131         test = "Satterthwaite", coefs = "All", p_values = TRUE)
4132
4133 # Retrieve variance decomposition
4134 life_phys_control_ICC_frame <- as.data.frame(VarCorr(life_phys_control))
4135 life_phys_control_ICC_frame
4136
4137 # Calculate R-squared compared to intercept-only model
4138 1 - sum(life_phys_control_ICC_frame[,4])/ sum(life_phys_intercept_ICC_frame[,4])
4139
4140 # ICC
4141 life_phys_control_ICC_frame[1, 4] / (life_phys_control_ICC_frame[1, 4] +
4142                                     life_phys_control_ICC_frame[2, 4])
4143
4144 # Removing year-fixed effects
4145 life_phys_rig_control_no_yearfixed <-
4146   lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism + lag_e_miinteco +
4147       lag_e_miinterc + lag_e_migdppcln + lag_e_migdpgro + lag_v2caviol +
4148       lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4149       data = datacomplete)
4150 coef_test(life_phys_rig_control_no_yearfixed, vcov = "CR2",
4151         test = "Satterthwaite", coefs = "All", p_values = TRUE)
4152
4153 # LR-test
4154
4155 # LR-test against personalism-only model
4156 deviance_life_phys_control <-
4157   deviance(lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) +
4158               latent_personalism + lag_e_miinteco + lag_e_miinterc +
4159               lag_e_migdppcln + lag_e_migdpgro + lag_v2caviol +
4160               lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4161               data = datacomplete, REML = FALSE))
4162 deviance_life_phys_personalism - deviance_life_phys_control
4163 pchisq(deviance_life_phys_personalism - deviance_life_phys_control,
4164        df = 9, lower.tail = FALSE)
4165
4166 # LR-test against model without year-fixed effects
4167 deviance_life_phys_control_no_yearfixed <-
4168   deviance(lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism +
4169               lag_e_miinteco + lag_e_miinterc + lag_e_migdppcln +
4170               lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4171               gwf_military + gwf_party, data = datacomplete, REML = FALSE))
4172 deviance_life_phys_control_no_yearfixed - deviance_life_phys_control
4173 pchisq(deviance_life_phys_control_no_yearfixed - deviance_life_phys_control,
4174        df = 54, lower.tail = FALSE)
4175
4176 # Previous model with rigour and impartiality of the public administration
4177 life_phys_rig_impact <-
4178   lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) + latent_personalism +
4179       v2clrspct + lag_e_miinteco + lag_e_miinterc + lag_e_migdppcln +
4180       lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4181       gwf_military + gwf_party, data = datacomplete)
4182
4183 # Calculate clustered standard errors, t-values, p-values and confidence
4184 # intervals
4185 coef_test(life_phys_rig_impact, vcov = "CR2", test = "Satterthwaite",
4186         coefs = "All", p_values = TRUE)
4187 conf_int(life_phys_rig_impact, vcov = "CR2", level = 0.9875,

```

```

4188     test = "Satterthwaite", coefs = "All", p_values = TRUE)
4189
4190 # Retrieve variance decomposition
4191 life_phys_rig_impact_ICC_frame <- as.data.frame(VarCorr(life_phys_rig_impact))
4192 life_phys_rig_impact_ICC_frame
4193
4194 # Calculate R-squared compared to intercept-only model
4195 1 - sum(life_phys_rig_impact_ICC_frame[,4])/
4196     sum(life_phys_intercept_ICC_frame[,4])
4197
4198 # ICC
4199 life_phys_rig_impact_ICC_frame[1, 4] / (life_phys_rig_impact_ICC_frame[1, 4] +
4200     life_phys_rig_impact_ICC_frame[2, 4])
4201
4202 # Removing year-fixed effects
4203 life_phys_rig_impact_no_yearfixed <-
4204     lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism + v2clrspct +
4205         lag_e_miinteco + lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4206         lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4207         data = datacomplete)
4208 coef_test(life_phys_rig_impact_no_yearfixed, vcov = "CR2",
4209     test = "Satterthwaite",
4210     coefs = "All", p_values = TRUE)
4211
4212 # LR-test
4213
4214 # LR-test against model without rigour and impartiality
4215 # of the public administration
4216 deviance_life_phys_rig_impact <-
4217     deviance(lmer(life_phys_x100 ~ (1|cowcode) + as.factor(year) +
4218         latent_personalism + v2clrspct + lag_e_miinteco +
4219         lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4220         lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military +
4221         gwf_party, data = datacomplete, REML = FALSE))
4222 deviance_life_phys_control - deviance_life_phys_rig_impact
4223 pchisq(deviance_life_phys_control - deviance_life_phys_rig_impact,
4224     df = 1, lower.tail = FALSE)
4225
4226 # LR-test against model without year-fixed effects
4227 deviance_life_phys_rig_impact_no_yearfixed <-
4228     deviance(lmer(life_phys_x100 ~ (1|cowcode) + latent_personalism + v2clrspct +
4229         lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
4230         lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4231         gwf_military + gwf_party, data = datacomplete, REML = FALSE))
4232 deviance_life_phys_rig_impact_no_yearfixed - deviance_life_phys_rig_impact
4233 pchisq(deviance_life_phys_rig_impact_no_yearfixed -
4234     deviance_life_phys_rig_impact, df = 54, lower.tail = FALSE)
4235
4236 # Residual diagnostics
4237
4238 # Add predicted values to dataset
4239 datacomplete$life_phys_predict <- predict(life_phys_rig_impact)
4240
4241 # Add residuals to dataset
4242
4243 # Calculate level 1 residuals
4244 datacomplete$life_phys_resid_level_1 <- residuals(life_phys_rig_impact)
4245
4246 # Standardise level 1 residuals

```

```

4247 for(c in unique(datacomplete$cowcode)) {
4248   datacomplete$life_phys_country_resid_sds[datacomplete$cowcode == c] <-
4249     sd(datacomplete$life_phys_resid_level_1[datacomplete$cowcode == c])
4250 }
4251 datacomplete$life_phys_stand_resid_level_1 <-
4252   datacomplete$life_phys_resid_level_1 /
4253   datacomplete$life_phys_country_resid_sds
4254
4255 # Calculate level 2 residuals
4256 life_phys_resid_level_2 <- as.data.frame(ranef(life_phys_rig_impact))
4257 life_phys_resid_level_2$cowcode <-
4258   as.numeric(levels(life_phys_resid_level_2$grp))[life_phys_resid_level_2$grp]
4259
4260 # Standardise level 2 residuals
4261 life_phys_resid_level_2$life_phys_stand_resid_level_2 <-
4262   life_phys_resid_level_2$condval / life_phys_rig_impact_ICC_frame$sdcor[1]
4263
4264 # Add level 2 residuals to dataset
4265 datacomplete <-
4266   left_join(datacomplete,
4267             life_phys_resid_level_2[, c("cowcode", "condval",
4268                                       "life_phys_stand_resid_level_2")],
4269             by = c("cowcode"))
4270 datacomplete <- rename(datacomplete, life_phys_resid_level_2 = condval)
4271
4272 # Autocorrelation
4273
4274 # Create dataframe with just the residuals
4275 life_phys_resid_pdata <-
4276   pdata.frame(datacomplete[, c("year", "cowcode", "life_phys_resid_level_1")],
4277               index = c("cowcode", "year"))
4278 # Create lagged residuals
4279 for (j in 1:20) {
4280   # Create the name of the lagged variable
4281   var <- paste("life_phys_lag", j, "_resid_level_1", sep = "")
4282   # Generate the lagged variable via plm's implementation of "lag"
4283   life_phys_resid_pdata[, var] <-
4284     plm::lag(life_phys_resid_pdata$life_phys_resid_level_1, k= j)
4285 }
4286
4287 # Calculate correlations between present values and lags
4288 autocors_life_phys_resid_level_1 <-
4289   cor(life_phys_resid_pdata$life_phys_resid_level_1,
4290       life_phys_resid_pdata[, 4:ncol(life_phys_resid_pdata)],
4291       use = "complete.obs")
4292
4293 # Reshape the correlations into a more workable format
4294 autocors_life_phys_resid_level_1 <-
4295   tibble(Lag = 1:20, Correlation = t(autocors_life_phys_resid_level_1))
4296
4297 # Create an autocorrelation plot
4298 ggplot(data = autocors_life_phys_resid_level_1, aes(x = Lag, y = Correlation)) +
4299   geom_col(fill = "lightblue") + blue_light
4300
4301 # Linearity
4302
4303 # Plot predicted values against level 1 residuals
4304 ggplot(datacomplete, aes(x = life_phys_predict,
4305                          y = life_phys_stand_resid_level_1)) +

```

```

4306 geom_point(colour = "lightblue") +
4307 geom_smooth(method = lm, formula = y ~ x, colour = "black",
4308             linetype = "dashed", se = FALSE) + blue_light +
4309 labs(x = "Predicted value", y = "Level 1 residual (standardised)")
4310
4311 # Plot values predicted by fixed part against level 2 residuals
4312 ggplot(datacomplete, aes(x = life_phys_predict - life_phys_resid_level_2,
4313                         y = life_phys_stand_resid_level_2)) +
4314 geom_point(colour = "lightblue") +
4315 geom_smooth(method = lm, formula = y ~ x, colour = "black",
4316             linetype = "dashed", se = FALSE) + blue_light +
4317 labs(x = "Fixed part predictions", y = "Level 2 residual (standardised)")
4318
4319 # Normality
4320
4321 # Assign residual levels to the residuals for combination
4322 rbind_resids_level_2 <-
4323   rename(life_phys_resid_level_2[, c("cowcode",
4324                                     "life_phys_stand_resid_level_2")],
4325          resid = life_phys_stand_resid_level_2)
4326 rbind_resids_level_2$level <- "Level 2"
4327 rbind_resids_level_1 <-
4328   rename(datacomplete[, c("cowcode", "life_phys_stand_resid_level_1")],
4329          resid = life_phys_stand_resid_level_1)
4330 rbind_resids_level_1$level <- "Level 1"
4331
4332 # Create long dataset of all residuals
4333 life_phys_resids <- rbind(rbind_resids_level_1, rbind_resids_level_2)
4334 rm(list = c("rbind_resids_level_1", "rbind_resids_level_2"))
4335
4336 # Plot residuals in dual QQ-plot
4337 withr::with_options(
4338   list(ggplot2.discrete.colour = c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
4339                                     "#0072B2", "#D55E00", "#CC79A7")),
4340   print(ggplot(data = life_phys_resids,
4341               aes(sample = resid, colour = as.factor(level))) + geom_qq() +
4342         geom_qq_line() + blue_light +
4343         scale_colour_discrete(name = "Residual level") +
4344         labs(y = "Standardised residuals",
4345              x = "Reference normal distribution")) )
4346
4347 # Plot residuals in single QQ-plots
4348
4349 # Level 1
4350 ggplot(data = life_phys_resids[life_phys_resids$level == "Level 1", ],
4351        aes(sample = resid)) +
4352   geom_qq(colour = "lightblue") + geom_qq_line() + blue_light +
4353   labs(y = "Standardised residuals", x = "Reference normal distribution")
4354
4355 # Level 2
4356 ggplot(data = life_phys_resids[life_phys_resids$level == "Level 2", ],
4357        aes(sample = resid)) +
4358   geom_qq(colour = "lightblue") + geom_qq_line() + blue_light +
4359   labs(y = "Standardised residuals", x = "Reference normal distribution")
4360
4361 ##### Freedom of movement #####
4362
4363 # Intercept-only model
4364 free_move_intercept <- lmer(freedom_movement ~ (1|cowcode), data = datacomplete)

```

```

4365
4366 # Calculate clustered standard errors, t-values, p-values and confidence
4367 # intervals
4368 coef_test(free_move_intercept, vcov = "CR2", test = "Satterthwaite",
4369           coefs = "All", p_values = TRUE)
4370 conf_int(free_move_intercept, vcov = "CR2", level = 0.9875,
4371          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4372
4373 # Retrieve variance decomposition
4374 free_move_intercept_ICC_frame <- as.data.frame(VarCorr(free_move_intercept))
4375 free_move_intercept_ICC_frame
4376
4377 # ICC
4378 free_move_intercept_ICC_frame[1, 4] / (free_move_intercept_ICC_frame[1, 4] +
4379                                       free_move_intercept_ICC_frame[2, 4])
4380
4381 # Calculate deviance
4382 deviance_free_move_intercept <- deviance(lmer(freedom_movement ~ (1|cowcode),
4383                                             data = datacomplete,
4384                                             REML = FALSE))
4385
4386 # Model with year-fixed effects
4387 free_move_yearfixed <- lmer(freedom_movement ~ (1|cowcode) + as.factor(year),
4388                           data = datacomplete)
4389
4390 # Calculate clustered standard errors, t-values, p-values and confidence
4391 # intervals
4392 coef_test(free_move_yearfixed, vcov = "CR2", test = "Satterthwaite",
4393           coefs = "All", p_values = TRUE)
4394 conf_int(free_move_yearfixed, vcov = "CR2", level = 0.9875,
4395          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4396
4397 # Retrieve variance decomposition
4398 free_move_yearfixed_ICC_frame <- as.data.frame(VarCorr(free_move_yearfixed))
4399 free_move_yearfixed_ICC_frame
4400
4401 # Calculate R-squared compared to intercept-only model
4402 1 - sum(free_move_yearfixed_ICC_frame[,4])/
4403     sum(free_move_intercept_ICC_frame[,4])
4404
4405 # ICC
4406 free_move_yearfixed_ICC_frame[1, 4] / (free_move_yearfixed_ICC_frame[1, 4] +
4407                                       free_move_yearfixed_ICC_frame[2, 4])
4408
4409 # LR-test
4410
4411 # LR-test against intercept-only model
4412 deviance_free_move_yearfixed <-
4413   deviance(lmer(freedom_movement ~ (1|cowcode) + as.factor(year),
4414               data = datacomplete, REML = FALSE))
4415 deviance_free_move_intercept - deviance_free_move_yearfixed
4416 pchisq(deviance_free_move_intercept - deviance_free_move_yearfixed,
4417        df = 54, lower.tail = FALSE)
4418
4419 # Model with personalism
4420 free_move_personalism <- lmer(freedom_movement ~ (1|cowcode) + as.factor(year) +
4421                             latent_personalism, data = datacomplete)
4422
4423 # Calculate clustered standard errors, t-values, p-values and confidence

```

```

4424 # intervals
4425 coef_test(free_move_personalism, vcov = "CR2", test = "Satterthwaite",
4426           coefs = "All", p_values = TRUE)
4427 conf_int(free_move_personalism, vcov = "CR2", level = 0.9875,
4428          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4429
4430 # Retrieve variance decomposition
4431 free_move_personalism_ICC_frame <- as.data.frame(VarCorr(free_move_personalism))
4432 free_move_personalism_ICC_frame
4433
4434 # Calculate R-squared compared to intercept-only model
4435 1 - sum(free_move_personalism_ICC_frame[,4])/
4436     sum(free_move_intercept_ICC_frame[,4])
4437
4438 # ICC
4439 free_move_personalism_ICC_frame[1, 4] /
4440 (free_move_personalism_ICC_frame[1, 4] +
4441  free_move_personalism_ICC_frame[2, 4])
4442
4443 # Removing year-fixed effects
4444 free_move_rig_personalism_no_yearfixed <-
4445   lmer(freedom_movement ~ (1|cowcode) + latent_personalism, data = datacomplete)
4446 coef_test(free_move_rig_personalism_no_yearfixed, vcov = "CR2",
4447           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4448
4449 # LR-test
4450
4451 # LR-test against model with year-fixed effects
4452 deviance_free_move_personalism <-
4453   deviance(lmer(freedom_movement ~ (1|cowcode) + as.factor(year) +
4454               latent_personalism, data = datacomplete, REML = FALSE))
4455
4456 deviance_free_move_yearfixed - deviance_free_move_personalism
4457 pchisq(deviance_free_move_yearfixed - deviance_free_move_personalism,
4458        df = 1, lower.tail = FALSE)
4459
4460 # LR-test against model without year-fixed effects
4461 deviance_free_move_personalism_no_yearfixed <-
4462   deviance(lmer(freedom_movement ~ (1|cowcode) + latent_personalism,
4463               data = datacomplete, REML = FALSE))
4464 deviance_free_move_personalism_no_yearfixed - deviance_free_move_personalism
4465 pchisq(deviance_free_move_personalism_no_yearfixed -
4466        deviance_free_move_personalism,
4467        df = 54, lower.tail = FALSE)
4468
4469 # Model with personalism and controls
4470 free_move_control <-
4471   lmer(freedom_movement ~ (1|cowcode) + as.factor(year) + latent_personalism +
4472       lag_e_miinteco + lag_e_miinterc + lag_e_migdppc1n + lag_e_migdpgro +
4473       lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4474       data = datacomplete)
4475
4476 # Calculate clustered standard errors, t-values, p-values and confidence
4477 # intervals
4478 coef_test(free_move_control, vcov = "CR2", test = "Satterthwaite",
4479           coefs = "All", p_values = TRUE)
4480 conf_int(free_move_control, vcov = "CR2", level = 0.9875,
4481          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4482

```

```

4483 # Retrieve variance decomposition
4484 free_move_control_ICC_frame <- as.data.frame(VarCorr(free_move_control))
4485 free_move_control_ICC_frame
4486
4487 # Calculate R-squared compared to intercept-only model
4488 1 - sum(free_move_control_ICC_frame[,4])/sum(free_move_intercept_ICC_frame[,4])
4489
4490 # ICC
4491 free_move_control_ICC_frame[1, 4] / (free_move_control_ICC_frame[1, 4] +
4492                                     free_move_control_ICC_frame[2, 4])
4493
4494 # Removing year-fixed effects
4495 free_move_rig_control_no_yearfixed <-
4496   lmer(freedom_movement ~ (1|cowcode) + latent_personalism + lag_e_miinteco +
4497       lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
4498       lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4499       data = datacomplete)
4500 coef_test(free_move_rig_control_no_yearfixed, vcov = "CR2",
4501           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4502
4503 # LR-test
4504
4505 # LR-test against personalism-only model
4506 deviance_free_move_control <-
4507   deviance(lmer(freedom_movement ~ (1|cowcode) + as.factor(year) +
4508               latent_personalism + lag_e_miinteco + lag_e_miinterc +
4509               lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
4510               lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4511               data = datacomplete, REML = FALSE))
4512 deviance_free_move_personalism - deviance_free_move_control
4513 pchisq(deviance_free_move_personalism - deviance_free_move_control,
4514        df = 9, lower.tail = FALSE)
4515
4516 # LR-test against model without year-fixed effects
4517 deviance_free_move_control_no_yearfixed <-
4518   deviance(lmer(freedom_movement ~ (1|cowcode) + latent_personalism +
4519               lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
4520               lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4521               gwf_military + gwf_party, data = datacomplete, REML = FALSE))
4522 deviance_free_move_control_no_yearfixed - deviance_free_move_control
4523 pchisq(deviance_free_move_control_no_yearfixed - deviance_free_move_control,
4524        df = 54, lower.tail = FALSE)
4525
4526 # Previous model with rigour and impartiality of the public administration
4527 free_move_rig_impact <-
4528   lmer(freedom_movement ~ (1|cowcode) + as.factor(year) + latent_personalism +
4529       v2clrspct + lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
4530       lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4531       gwf_military + gwf_party, data = datacomplete)
4532
4533 # Calculate clustered standard errors, t-values, p-values and confidence
4534 # intervals
4535 coef_test(free_move_rig_impact, vcov = "CR2", test = "Satterthwaite",
4536           coefs = "All", p_values = TRUE)
4537 conf_int(free_move_rig_impact, vcov = "CR2", level = 0.9875,
4538          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4539
4540 # Retrieve variance decomposition
4541 free_move_rig_impact_ICC_frame <- as.data.frame(VarCorr(free_move_rig_impact))

```

```

4542 free_move_rig_impact_ICC_frame
4543
4544 # Calculate R-squared compared to intercept-only model
4545 1 - sum(free_move_rig_impact_ICC_frame[,4])/
4546     sum(free_move_intercept_ICC_frame[,4])
4547
4548 # ICC
4549 free_move_rig_impact_ICC_frame[1, 4] / (free_move_rig_impact_ICC_frame[1, 4] +
4550     free_move_rig_impact_ICC_frame[2, 4])
4551
4552 # Removing year-fixed effects
4553 free_move_rig_impact_no_yearfixed <-
4554   lmer(freedom_movement ~ (1|cowcode) + latent_personalism + v2clrspct +
4555     lag_e_miinteco + lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4556     lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4557     data = datacomplete)
4558 coef_test(free_move_rig_impact_no_yearfixed, vcov = "CR2",
4559     test = "Satterthwaite", coefs = "All", p_values = TRUE)
4560
4561 # LR-test
4562
4563 # LR-test against model without rigour and impartiality
4564 # of the public administration
4565 deviance_free_move_rig_impact <-
4566   deviance(lmer(freedom_movement ~ (1|cowcode) + as.factor(year) +
4567     latent_personalism + v2clrspct + lag_e_miinteco +
4568     lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4569     lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military +
4570     gwf_party, data = datacomplete, REML = FALSE))
4571 deviance_free_move_control - deviance_free_move_rig_impact
4572 pchisq(deviance_free_move_control - deviance_free_move_rig_impact,
4573     df = 1, lower.tail = FALSE)
4574
4575 # LR-test against model without year-fixed effects
4576 deviance_free_move_rig_impact_no_yearfixed <-
4577   deviance(lmer(freedom_movement ~ (1|cowcode) + latent_personalism +
4578     v2clrspct + lag_e_miinteco + lag_e_miinterc +
4579     lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
4580     lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4581     data = datacomplete, REML = FALSE))
4582 deviance_free_move_rig_impact_no_yearfixed - deviance_free_move_rig_impact
4583 pchisq(deviance_free_move_rig_impact_no_yearfixed -
4584     deviance_free_move_rig_impact, df = 54, lower.tail = FALSE)
4585
4586 # Residual diagnostics
4587
4588 # Add predicted values to dataset
4589 datacomplete$free_move_predict <- predict(free_move_rig_impact)
4590
4591 # Add residuals to dataset
4592
4593 # Calculate level 1 residuals
4594 datacomplete$free_move_resid_level_1 <- residuals(free_move_rig_impact)
4595
4596 # Standardise level 1 residuals
4597 for(c in unique(datacomplete$cowcode)) {
4598   datacomplete$free_move_country_resid_sds[datacomplete$cowcode == c] <-
4599     sd(datacomplete$free_move_resid_level_1[datacomplete$cowcode == c])
4600 }

```



```

4601 datacomplete$free_move_stand_resid_level_1 <-
4602   datacomplete$free_move_resid_level_1 /
4603   datacomplete$free_move_country_resid_sds
4604
4605 # Calculate level 2 residuals
4606 free_move_resid_level_2 <- as.data.frame(ranef(free_move_rig_impact))
4607 free_move_resid_level_2$cowcode <-
4608   as.numeric(levels(free_move_resid_level_2$grp))[free_move_resid_level_2$grp]
4609
4610 # Standardise level 2 residuals
4611 free_move_resid_level_2$free_move_stand_resid_level_2 <-
4612   free_move_resid_level_2$condval / free_move_rig_impact_ICC_frame$sdcor[1]
4613
4614 # Add level 2 residuals to dataset
4615 datacomplete <-
4616   left_join(datacomplete,
4617             free_move_resid_level_2[, c("cowcode", "condval",
4618                                       "free_move_stand_resid_level_2")],
4619             by = c("cowcode"))
4620 datacomplete <- rename(datacomplete, free_move_resid_level_2 = condval)
4621
4622 # Autocorrelation
4623
4624 # Create dataframe with just the residuals
4625 free_move_resid_pdata <-
4626   pdata.frame(datacomplete[, c("year", "cowcode", "free_move_resid_level_1")],
4627               index = c("cowcode", "year"))
4628 # Create lagged residuals
4629 for (j in 1:20) {
4630   # Create the name of the lagged variable
4631   var <- paste("free_move_lag", j, "_resid_level_1", sep = "")
4632   # Generate the lagged variable via plm's implementation of "lag"
4633   free_move_resid_pdata[, var] <-
4634     plm::lag(free_move_resid_pdata$free_move_resid_level_1, k= j)
4635 }
4636
4637 # Calculate correlations between present values and lags
4638 autocors_free_move_resid_level_1 <-
4639   cor(free_move_resid_pdata$free_move_resid_level_1,
4640       free_move_resid_pdata[, 4:ncol(free_move_resid_pdata)],
4641       use = "complete.obs")
4642
4643 # Reshape the correlations into a more workable format
4644 autocors_free_move_resid_level_1 <-
4645   tibble(Lag = 1:20, Correlation = t(autocors_free_move_resid_level_1))
4646
4647 # Create an autocorrelation plot
4648 ggplot(data = autocors_free_move_resid_level_1, aes(x = Lag, y = Correlation)) +
4649   geom_col(fill = "lightblue") + blue_light
4650
4651 # Linearity
4652
4653 # Plot predicted values against level 1 residuals
4654 ggplot(datacomplete, aes(x = free_move_predict,
4655                          y = free_move_stand_resid_level_1)) +
4656   geom_point(colour = "lightblue") +
4657   geom_smooth(method = lm, formula = y ~ x, colour = "black",
4658              linetype = "dashed", se = FALSE) + blue_light +
4659   labs(x = "Predicted value", y = "Level 1 residual (standardised)")

```

```

4660
4661 #       Plot fixed part predictions against level 2 residuals
4662 ggplot(datacomplete, aes(x = free_move_predict - free_move_resid_level_2,
4663                          y = free_move_stand_resid_level_2)) +
4664   geom_point(colour = "lightblue") +
4665   geom_smooth(method = lm, formula = y ~ x, colour = "black",
4666              linetype = "dashed", se = FALSE) + blue_light +
4667   labs(x = "Fixed part predictions", y = "Level 2 residual (standardised)")
4668
4669 #       Normality
4670
4671 #       Assign residual levels to the residuals for combination
4672 rbind_resids_level_2 <-
4673   rename(free_move_resid_level_2[, c("cowcode",
4674                                     "free_move_stand_resid_level_2")],
4675          resid = free_move_stand_resid_level_2)
4676 rbind_resids_level_2$level <- "Level 2"
4677 rbind_resids_level_1 <-
4678   rename(datacomplete[, c("cowcode", "free_move_stand_resid_level_1")],
4679          resid = free_move_stand_resid_level_1)
4680 rbind_resids_level_1$level <- "Level 1"
4681
4682 #       Create long dataset of all residuals
4683 free_move_resids <- rbind(rbind_resids_level_1, rbind_resids_level_2)
4684 rm(list = c("rbind_resids_level_1", "rbind_resids_level_2"))
4685
4686 #       Plot residuals in dual QQ-plot
4687 withr::with_options(
4688   list(ggplot2.discrete.colour = c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
4689                                   "#0072B2", "#D55E00", "#CC79A7")),
4690   print(ggplot(data = free_move_resids, aes(sample = resid,
4691                                             colour = as.factor(level))) +
4692         geom_qq() + geom_qq_line() + blue_light +
4693         scale_colour_discrete(name = "Residual level") +
4694         labs(y = "Standardised residuals",
4695              x = "Reference normal distribution")) )
4696
4697 #       Plot residuals in single QQ-plots
4698 #       Level 1
4699 ggplot(data = free_move_resids[free_move_resids$level == "Level 1", ],
4700        aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
4701   blue_light + labs(y = "Standardised residuals",
4702                    x = "Reference normal distribution")
4703
4704 #       Level 2
4705 ggplot(data = free_move_resids[free_move_resids$level == "Level 2", ],
4706        aes(sample = resid)) + geom_qq(colour = "lightblue") + geom_qq_line() +
4707   blue_light + labs(y = "Standardised residuals",
4708                    x = "Reference normal distribution")
4709
4710 ##### Rigour and impartiality of the public administration #####
4711
4712 #       Random intercept-only model
4713 rig_impact_intercept <- lmer(v2clrspct ~ (1|cowcode), data = datacomplete)
4714
4715 #       Calculate clustered standard errors, t-values, p-values and confidence
4716 #       intervals
4717 coef_test(rig_impact_intercept, vcov = "CR2", test = "Satterthwaite",
4718          coefs = "All", p_values = TRUE)

```

```

4719 conf_int(rig_impact_intercept, vcov = "CR2", level = 0.9875,
4720           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4721
4722 # Retrieve variance decomposition
4723 rig_impact_intercept_ICC_frame <- as.data.frame(VarCorr(rig_impact_intercept))
4724 rig_impact_intercept_ICC_frame
4725
4726 # ICC
4727 rig_impact_intercept_ICC_frame[1, 4] / (rig_impact_intercept_ICC_frame[1, 4] +
4728                                         rig_impact_intercept_ICC_frame[2, 4])
4729
4730 # Calculate deviance
4731 deviance_rig_impact_intercept <- deviance(lmer(v2clrspct ~ (1|cowcode),
4732                                             data = datacomplete,
4733                                             REML = FALSE))
4734
4735 # Model with time-fixed effects
4736 rig_impact_yearfixed <- lmer(v2clrspct ~ (1|cowcode) + as.factor(year),
4737                             data = datacomplete)
4738
4739 # Calculate clustered standard errors, t-values, p-values and confidence
4740 # intervals
4741 coef_test(rig_impact_yearfixed, vcov = "CR2", test = "Satterthwaite",
4742           coefs = "All", p_values = TRUE)
4743 conf_int(rig_impact_yearfixed, vcov = "CR2", level = 0.9875,
4744           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4745
4746 # Retrieve variance decomposition
4747 rig_impact_yearfixed_ICC_frame <- as.data.frame(VarCorr(rig_impact_yearfixed))
4748 rig_impact_yearfixed_ICC_frame
4749
4750 # Calculate R-squared compared to intercept-only model
4751 1 - sum(rig_impact_yearfixed_ICC_frame[,4])/
4752       sum(rig_impact_intercept_ICC_frame[,4])
4753
4754 # ICC
4755 rig_impact_yearfixed_ICC_frame[1, 4] / (rig_impact_yearfixed_ICC_frame[1, 4] +
4756                                         rig_impact_yearfixed_ICC_frame[2, 4])
4757
4758 # LR-test against intercept-only model
4759 deviance_rig_impact_yearfixed <-
4760   deviance(lmer(v2clrspct ~ (1|cowcode) + as.factor(year), data = datacomplete,
4761               REML = FALSE))
4762 deviance_rig_impact_intercept - deviance_rig_impact_yearfixed
4763 pchisq(deviance_rig_impact_intercept - deviance_rig_impact_yearfixed,
4764         df = 54, lower.tail = FALSE)
4765
4766 # Model with personalism
4767 rig_impact_personalism <- lmer(v2clrspct ~ (1|cowcode) + as.factor(year) +
4768                               latent_personalism, data = datacomplete)
4769
4770 # Calculate clustered standard errors, t-values, p-values and confidence
4771 # intervals
4772 coef_test(rig_impact_personalism, vcov = "CR2", test = "Satterthwaite",
4773           coefs = "All", p_values = TRUE)
4774 conf_int(rig_impact_personalism, vcov = "CR2", level = 0.9875,
4775           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4776
4777 # Retrieve variance decomposition

```

```

4778 rig_impact_personalism_ICC_frame <-
4779   as.data.frame(VarCorr(rig_impact_personalism))
4780 rig_impact_personalism_ICC_frame
4781
4782 #       Calculate R-squared compared to intercept-only model
4783 1 - sum(rig_impact_personalism_ICC_frame[,4])/
4784   sum(rig_impact_intercept_ICC_frame[,4])
4785
4786 #       ICC
4787 rig_impact_personalism_ICC_frame[1, 4] /
4788   (rig_impact_personalism_ICC_frame[1, 4] +
4789     rig_impact_personalism_ICC_frame[2, 4])
4790
4791 #       Removing year-fixed effects
4792 rig_impact_personalism_no_yearfixed <-
4793   lmer(v2clrspct ~ (1|cowcode) + latent_personalism, data = datacomplete)
4794 coef_test(rig_impact_personalism_no_yearfixed, vcov = "CR2",
4795           test = "Satterthwaite", coefs = "All", p_values = TRUE)
4796
4797 #       LR-test
4798
4799 #       LR-test against year-fixed effects model
4800 deviance_rig_impact_personalism <-
4801   deviance(lmer(v2clrspct ~ (1|cowcode) + as.factor(year) + latent_personalism,
4802               data = datacomplete, REML = FALSE))
4803 deviance_rig_impact_yearfixed - deviance_rig_impact_personalism
4804 pchisq(deviance_rig_impact_intercept - deviance_rig_impact_personalism,
4805         df = 1, lower.tail = FALSE)
4806
4807 #       LR-test against model without year-fixed effects
4808 deviance_rig_impact_personalism_no_yearfixed <-
4809   deviance(lmer(v2clrspct ~ (1|cowcode) + latent_personalism,
4810               data = datacomplete, REML = FALSE))
4811 deviance_rig_impact_personalism_no_yearfixed - deviance_rig_impact_personalism
4812 pchisq(deviance_rig_impact_personalism_no_yearfixed -
4813         deviance_rig_impact_personalism,
4814         df = 54, lower.tail = FALSE)
4815
4816 #       Model with personalism and controls
4817 rig_impact_control <-
4818   lmer(v2clrspct ~ (1|cowcode) + as.factor(year) + latent_personalism +
4819         lag_e_miinteco + lag_e_miinterc + lag_e_migdppc1n + lag_e_migdpgro +
4820         lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4821         data = datacomplete)
4822
4823 #       Calculate clustered standard errors, t-values, p-values and confidence
4824 #       intervals
4825 coef_test(rig_impact_control, vcov = "CR2", test = "Satterthwaite",
4826           coefs = "All", p_values = TRUE)
4827 conf_int(rig_impact_control, vcov = "CR2", level = 0.9875,
4828          test = "Satterthwaite", coefs = "All", p_values = TRUE)
4829
4830 #       Retrieve variance decomposition
4831 rig_impact_control_ICC_frame <- as.data.frame(VarCorr(rig_impact_control))
4832 rig_impact_control_ICC_frame
4833
4834 #       Calculate R-squared compared to intercept-only model
4835 1 - sum(rig_impact_control_ICC_frame[,4])/
4836   sum(rig_impact_intercept_ICC_frame[,4])

```

```

4837
4838 #       ICC
4839 rig_impact_control_ICC_frame[1, 4] /
4840   (rig_impact_control_ICC_frame[1, 4] +
4841     rig_impact_control_ICC_frame[2, 4])
4842
4843 #       Removing year-fixed effects
4844 rig_impact_control_no_yearfixed <-
4845   lmer(v2clrspct ~ (1|cowcode) + latent_personalism + lag_e_miinteco +
4846     lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro + lag_v2caviol +
4847     lag_log10pop + gwf_monarch + gwf_military + gwf_party,
4848     data = datacomplete)
4849 coef_test(rig_impact_control_no_yearfixed, vcov = "CR2", test = "Satterthwaite",
4850   coefs = "All", p_values = TRUE)
4851
4852 #       LR-test
4853
4854 #       LR-test against model without regime type controls
4855 deviance_rig_impact_control <-
4856   deviance(lmer(v2clrspct ~ (1|cowcode) + as.factor(year) + latent_personalism +
4857     lag_e_miinteco + lag_e_miinterc + lag_e_migdppln +
4858     lag_e_migdpgro + lag_v2caviol + lag_log10pop + gwf_monarch +
4859     gwf_military + gwf_party, data = datacomplete, REML = FALSE))
4860 deviance_rig_impact_personalism - deviance_rig_impact_control
4861 pchisq(deviance_rig_impact_personalism - deviance_rig_impact_control,
4862   df = 9, lower.tail = FALSE)
4863
4864 #       LR-test against model without year-fixed effects
4865 deviance_rig_impact_control_no_yearfixed <-
4866   deviance(lmer(v2clrspct ~ (1|cowcode) + latent_personalism + lag_e_miinteco +
4867     lag_e_miinterc + lag_e_migdppln + lag_e_migdpgro +
4868     lag_v2caviol + lag_log10pop + gwf_monarch + gwf_military +
4869     gwf_party, data = datacomplete, REML = FALSE))
4870 deviance_rig_impact_control_no_yearfixed - deviance_rig_impact_control
4871 pchisq(deviance_rig_impact_control_no_yearfixed - deviance_rig_impact_control,
4872   df = 54, lower.tail = FALSE)
4873
4874 #       Residual diagnostics
4875
4876 #       Add predicted values to dataset
4877 datacomplete$rig_impact_predict <- predict(rig_impact_control)
4878
4879 #       Add residuals to dataset
4880
4881 #       Calculate level 1 residuals
4882 datacomplete$rig_impact_resid_level_1 <- residuals(rig_impact_control)
4883
4884 #       Standardise level 1 residuals
4885 for(c in unique(datacomplete$cowcode)) {
4886   datacomplete$rig_impact_country_resid_sds[datacomplete$cowcode == c] <-
4887     sd(datacomplete$rig_impact_resid_level_1[datacomplete$cowcode == c])
4888 }
4889 datacomplete$rig_impact_stand_resid_level_1 <-
4890   datacomplete$rig_impact_resid_level_1 /
4891   datacomplete$rig_impact_country_resid_sds
4892
4893 #       Calculate level 2 residuals
4894 rig_impact_resid_level_2 <- as.data.frame(ranef(rig_impact_control))
4895 rig_impact_resid_level_2$cowcode <-

```

```

4896     as.numeric(levels(rig_impact_resid_level_2$grp))[rig_impact_resid_level_2$grp]
4897
4898 #     Standardise level 2 residuals
4899 rig_impact_resid_level_2$rig_impact_stand_resid_level_2 <-
4900     rig_impact_resid_level_2$condval / rig_impact_control_ICC_frame$sdcor[1]
4901
4902 #     Add level 2 residuals to dataset
4903 datacomplete <-
4904     left_join(datacomplete,
4905             rig_impact_resid_level_2[, c("cowcode", "condval",
4906                                         "rig_impact_stand_resid_level_2")],
4907             by = c("cowcode"))
4908 datacomplete <- rename(datacomplete, rig_impact_resid_level_2 = condval)
4909
4910 #     Autocorrelation
4911
4912 #     Create dataframe with just the residuals
4913 rig_impact_resid_pdata <-
4914     pdata.frame(datacomplete[, c("year", "cowcode", "rig_impact_resid_level_1")],
4915               index = c("cowcode", "year"))
4916 #     Create lagged residuals
4917 for (j in 1:20) {
4918     # Create the name of the lagged variable
4919     var <- paste("rig_impact_lag", j, "_resid_level_1", sep = "")
4920     # Generate the lagged variable via plm's implementation of "lag"
4921     rig_impact_resid_pdata[, var] <-
4922     plm::lag(rig_impact_resid_pdata$rig_impact_resid_level_1, k= j)
4923 }
4924
4925 #     Calculate correlations between present values and lags
4926 autocors_rig_impact_resid_level_1 <-
4927     cor(rig_impact_resid_pdata$rig_impact_resid_level_1,
4928         rig_impact_resid_pdata[, 4:ncol(rig_impact_resid_pdata)],
4929         use = "complete.obs")
4930
4931 #     Reshape the correlations into a more workable format
4932 autocors_rig_impact_resid_level_1 <-
4933     tibble(Lag = 1:20, Correlation = t(autocors_rig_impact_resid_level_1))
4934
4935 #     Create an autocorrelation plot
4936 ggplot(data = autocors_rig_impact_resid_level_1, aes(x = Lag,
4937                                                    y = Correlation)) +
4938     geom_col(fill = "lightblue") + blue_light
4939
4940 #     Linearity
4941
4942 #     Plot predicted values against level 1 residuals
4943 ggplot(datacomplete, aes(x = rig_impact_predict,
4944                         y = rig_impact_stand_resid_level_1)) +
4945     geom_point(colour = "lightblue") +
4946     geom_smooth(method = lm, formula = y ~ x, colour = "black",
4947               linetype = "dashed", se = FALSE) + blue_light +
4948     labs(x = "Predicted value", y = "Level 1 residual (standardised)")
4949
4950 #     Plot fixed part predictions against level 2 residuals
4951 ggplot(datacomplete, aes(x = rig_impact_predict - rig_impact_resid_level_2,
4952                         y = rig_impact_stand_resid_level_2)) +
4953     geom_point(colour = "lightblue") +
4954     geom_smooth(method = lm, formula = y ~ x, colour = "black",

```

```

4955         linetype = "dashed", se = FALSE) + blue_light +
4956     labs(x = "Fixed part predictions", y = "Level 2 residual (standardised)")
4957
4958 #     Normality
4959
4960 #     Assign residual levels to the residuals for combination
4961 rbind_resids_level_2 <-
4962     rename(rig_impact_resid_level_2[, c("cowcode",
4963         "rig_impact_stand_resid_level_2")],
4964         resid = rig_impact_stand_resid_level_2)
4965 rbind_resids_level_2$level <- "Level 2"
4966 rbind_resids_level_1 <-
4967     rename(datacomplete[, c("cowcode", "rig_impact_stand_resid_level_1")],
4968         resid = rig_impact_stand_resid_level_1)
4969 rbind_resids_level_1$level <- "Level 1"
4970
4971 #     Create long dataset of all residuals
4972 rig_impact_resids <- rbind(rbind_resids_level_1, rbind_resids_level_2)
4973 rm(list = c("rbind_resids_level_1", "rbind_resids_level_2"))
4974
4975 #     Plot residuals in dual QQ-plot
4976 withr::with_options(
4977     list(ggplot2.discrete.colour = c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
4978         "#0072B2", "#D55E00", "#CC79A7")),
4979     print(ggplot(data = rig_impact_resids, aes(sample = resid,
4980         colour = as.factor(level))) +
4981         geom_qq() + geom_qq_line() + blue_light +
4982         scale_colour_discrete(name = "Residual level") +
4983         labs(y = "Standardised residuals",
4984             x = "Reference normal distribution")) )
4985
4986 #     Plot residuals in single QQ-plots
4987
4988 #     Level 1
4989 ggplot(data = rig_impact_resids[rig_impact_resids$level == "Level 1", ],
4990     aes(sample = resid)) + geom_qq(colour = "lightblue") +
4991     geom_qq_line() + blue_light + labs(y = "Standardised residuals",
4992         x = "Reference normal distribution")
4993
4994 #     Level 2
4995 ggplot(data = rig_impact_resids[rig_impact_resids$level == "Level 2", ],
4996     aes(sample = resid)) + geom_qq(colour = "lightblue") +
4997     geom_qq_line() + blue_light + labs(y = "Standardised residuals",
4998         x = "Reference normal distribution")

```