To comply or not to comply?

Building an index for constitutional compliance

Thea Christine Dybvig

Masteroppgave

Spring 2021

Department of Comparative Politics
University of Bergen

Abstract

There are limited, but increasing, attempts at quantitatively measuring the distance between de jure provisions and their de facto adherence. Such attempts seem mostly concerned with rights provisions, but some also concern the independence of the judiciary. This thesis seeks to enter the discussion on how such measurement should happen by contributing an index of constitutional compliance for a set of 175 states through 41 years. The index attempts to measure both structural and rights provisions. It is divided into four components, one for provisions regarding the executive, one for the legislature, one for the judiciary and public administration, and one for rights.

The index is further analysed to assess amongst predictors from the literature on state repression of de facto rights, what might prove to be influencing a broader concept of constitutional compliance. The findings indicate that executive compliance with the high court's decisions and electoral democracy lead to increases in compliance scores across countries. The comprehensiveness of the constitution, or how many provisions are present, have a curvilinear relationship with compliance but also seems a consistent predictor. The main determinant of compliance seems to be, however, path dependency.

Constitutions perform many functions, where constitutionalism is of particular interest for this thesis. Constitutional compliance is here conceptualised as the mechanism through which restraining and enabling of government power happens. The index is a first step towards further operationalising the concept of constitutionalism.

Aknowledgements

Først og fremst ønsker jeg å takke veilederen min, Jonas Linde. Du har vært utrolig tålmodig med meg og alltid vært villig til å stille opp og hjelpe, noe som har betydd mye i en veldig kronglete prosess.

Deretter ønsker jeg å takke alle mine medstudenter. Dere har vært helt fantastiske, og lesesaler har ikke noe med å være så trivelig som våre var. Jeg savner allerede all den utrolig dårlige pensumhumoren, rivaliseringen over absolutt alt trivielt det kunne rivaliseres over og det gode samholdet. Det var slitsomt, men gøy, og jeg kommer sikkert til å påstå at jeg til og med savner de sene kveldene. Etterhvert.

Til slutt ønsker jeg å takke familien min, som har stått for mye omsorg og drahjelp. Jeg føler meg utrolig heldig som har dere, og ingenting hadde gått uten dere.

Tusen takk til alle sammen ☺

Table of Contents

List of tables

List of figures

1 Introduction

"Why would the government behave like the proverbial chicken that stays inside the chalk circle it could easily transgress?" (Elster 2015, 442).

There seems to be some sort of perpetual paradox connected to the authority of constitutions. "Rather than being presented as an exchange of promises between separate parties, modern constitutions are typically styled as frameworks which "we the people" give ourselves (Holmes 1988, 209). What power constitutions could possibly hold have been a philosophical as well as empirically grounded concern. If a sovereign is, well, sovereign, how can he bind himself? He is powerless if he could be bound by the past, but similarly powerless if he cannot bind himself for the future (Hardin 1999, 162). The argument, however, changes when the agent making the promise to itself is not a single agent. A state consists of many actors with interests in the continued functioning of the institutions that make up government, which is how compliance is generally explained (Tushnet 2014, 37).

While there are many theoretical explanations for constitutional compliance, there are few attempts at measuring the phenomenon systematically across countries. "Still, a document's bearing the label "a constitution" and declaring its own control over all other political acts proves nothing. We need to distinguish between the authority a text asserts and the authority it exerts" (Murphy 1993, 7-8). This is precisely the aim of this thesis; to distinguish between de jure provisions and de facto adherence to them.

> "Yet the category of sham constitutions is inevitably imperfect. Practice in almost every nation will fail to correspond with some aspects of that nation's formal constitution, at least from some perspective, and so we need a metric for determining when the shortfall is great enough to make the constitution a sham. That metric is again almost inevitably going to be a matter of controversy: How much weight should it give to shortfalls with respect to rights as against shortfalls with respect to government structure, for example?" (Tushnet 2014, 11)

This endeavour is not altogether straightforward, however, as measuring adherence to a constitution is both empirically and theoretically challenging. Constitutions can be rather long and complex documents with some rather specific provisions. They frequently include "excruciating detail" on flags, anthems and other national symbols, or specifies currencies without taking inflation into account (Elkins, Ginsburg, Melton 2009, 52). Not all provisions can then be of interest for a compliance measure. The second question that arises is then *compliance for what reason?* In this thesis compliance is understood as a mechanism through which the constitution can perform different functions. It is more specifically operationalised for the performance of constitutionalism.

1.1 Structure of the thesis

The aim of this thesis is to explore the possibility of operationalising and measuring constitutionalism. Measuring constitutionalism itself seems too big a task for a project of such scope as this, and the focus is rather on measurement that would be necessary for the operationalisation of constitutionalism. By this, I mean that measuring constitutional compliance is introduced as a starting point for operationalisation of constitutionalism.

My research question is therefore *how does one measure constitutional compliance? And what can we expect to influence levels of constitutional adherence?*

The structure of this thesis is then as follows. Chapter 2 provides an overview of the theoretical foundations of the thesis. Literature on constitutional functions in general, and more specifically, constitutionalism is introduced, before other quantitative research into constitutional compliance is discussed. In Chapter 3, the operationalisation of constitutionalism and the methodological framework for constructing the compliance index is presented. Chapter 4 provides an overview of cursory findings from the index. Chapter 5 presents a more in-depth attempt at analysis of the compliance measure using multilevel modelling to assess predictors of constitutional compliance. Lastly, Chapter 6 provides the concluding remarks.

1.2 Why study constitutionalism and compliance?

In order to assess the roles constitutions play in different contexts, it is important to operationalise and measure the different constitutional functions. Measuring constitutional compliance is an attempt at moving towards operationalisation of constitutionalism as a function of the constitution. The argument put forth by this thesis is that compliance is the mechanism through which constitutionalism happens. There are many different conceptualisations of constitutionalism, and while I operate with one myself, the central argument and therefore motivation of the thesis is that regardless of what attributes are included in the theoretical definition of constitutionalism, constitutional compliance is necessary. As such, the ability to classify cases of constitutional compliance and non-compliance becomes integral to any attempted empirical assessment of constitutionalism. My aim is therefore to engage in the question of how this could be done.

1.3 How to study constitutional compliance?

The empirical study of constitutional compliance ought, in my opinion, be subject to both qualitative and quantitative inquiries. If qualitative work can foster questions about how things might be connected at a broader level and "[l]ikewise, when statistical results about the effects of causes are reported, it seems natural to ask if these results make sense in terms of the history of individual cases; one wishes to try to locate the effects in specific cases" (Mahoney and Goertz 2006, 231). The measure here is created with this hope in mind.

Attempting to describe the gap between de jure and de facto by using snapshots in time disallows the effect of time, which will then bias the findings. Constitutions often involve describing institutions as they should be and allowing for the possibility of a maturing effect in their study is therefore important (Elkins, Ginsburg, and Melton 2016, 236). This does not mean disregarding any gaps with the excuse of aspirations, but rather allowing that any snapshot in time will inevitably capture states at different points in their constitutional lifespan.

There are definitive drawbacks to studying constitutions quantitatively, however. To see if it is even possible to draw any inferences about constitutional compliance it is necessary to study it across historical, economic, cultural, legal, political, and social circumstances, which all ought to be taken into account. Since the methods applied here are quantitative only, these contextual differences escape measurement. That is not to say any quantitative venture into the

world of constitutions is useless or limited in its contributions. Different tools have different purposes, and a larger toolkit ought to be more suited to solving more issues.

1.4 Findings

The thesis' main result is in many ways the creation of a compliance index. The index encompasses 175 countries and a time period of 41 years, from 1980 to 2020. A central aim of this endeavour is to further research into and discussions around how to measure concepts such as compliance, and more broadly constitutionalism, though that is too broad of a concept to deal with fully here. The compliance index is still a work in progress, as data on more provisions ought to be added for a more complete measure. Chapter 4 provides an overview into the index, with some numbers on the variables generated and a cursory look into some bivariate correlations. Chapter 5 provides a more in-depth analysis, through a multilevel regression, and findings seem to be that electoral democracy, executive compliance with the decisions of the high court on decisions it does not favour, and a quadratic term for the number of provisions present all affect levels of compliance. The first two relationships are positive, and the latter of a convex nature.

2 Constitutions and constitutionalism

> "[…] Nowadays "constitution" has become an ambiguous term, covering two very different meanings: a strict, substantive meaning (the *garantiste* meaning), and a formal, cosmic meaning. It follows from this that whereas in the 19th century a question such as, "What is the role of a constitution in a political system?", could be answered without asking first, "What is a constitution?", this is no longer the case." (Sartori 1962, 857).

Sartori (1962, 857) writes of how the answer to this first question becomes banal and uninteresting if a constitution is defined narrowly as an organizational "map" of exercisable power in the polity, but also how the answer is different for each country one would examine. The first part of this position has not been shared by all. Later academics[1] have expanded greatly on what roles a constitution can play in a political system, even if such a system is not a liberal democratic one capable of the *garantiste* function. Some have even done work on how they perform in autocratic regimes.

The second part of Sartori's position, however, seems commonly accepted given the tendency to qualitative methods in the fields of comparative law and constitutions. Such a "Montesquieuan" approach to the distinctiveness of national constitutions suggests there can be limited means for comparison across differing countries (Tushnet 2014, 2-4). There have, however, been efforts to identify "[…] functions common to all constitutional systems […]", following a more "Benthamite" logic of functionalism expressed as a "[…] universal grammar of government" (Tushnet 2014, 3). There have also been efforts to consider what roles a constitution might play in a theoretical entity and to identify where such might have occurred (see for instance Galligan and Versteeg 2013).

The following chapter starts with considerations of several functions constitutions might have in different political, social and historical systems, before focusing more specifically on constitutionalism. This last function is at the core of what Sartori (1962, 859-860) argued a constitution ought to do, or what its purpose should be understood as. Constitutionalism is certainly established today as one of, if not the, most important effects of the constitutional

[1] See for instance Ginsburg and Simpser 2014

order (Elkins, Ginsburg, Melton 2009, 38). Following is a discussion on concepts of constitutionalism and more specifically a discussion on compliance as a necessary component of such. In the last part of the chapter, focus turns to previous, quantitative research on the effects of de jure provisions and their corresponding, de facto situations.

2.1 Defining constitutions

Constitution has sometimes been thought to be a somewhat ambiguous term, though differing understandings can typically be separated into two categories defined either in terms of their form or their functions. The word can be understood to refer to the wider constitutional order or more specifically to a written text often titled "The Constitution" (Elkins, Ginsburg, and Melton 2009, 38-39). The constitutional order comprises of the many functions a constitution might perform for its constituency, through the text itself or through other texts or nontextual instruments (Elkins, Ginsburg and Melton 2009, 40-47).

Understanding the constitution through its functions means understanding the constitution as social and political phenomena. They are at once independent documents outlining the rules of government while also being interdependent upon their political context, in which they are constantly interpreted and oftentimes amended (Galligan and Versteeg 2013, 7). This interplay means that a constitution might serve several differing functions simultaneously, while also performing different functions at different times given evolving social and political circumstances. This section seeks to explore some of these functions.

2.1.1 Expressions of value

One function of the constitution is its symbolic power; constitutions can be expressions of value (Elkins, Melton and Ginsburg 2009, 38; Galligan and Versteeg 2013, 8-14).
These values can be national ones, reflecting a sense of shared history and identity, and they can be transnational; disseminated through outside coercion, competition between states, learning or a desire for international legitimacy and acceptance (Galligan and Versteeg 2013, 15-18). Whether these values actually influence the *behaviour* of government officials and institutions boils down to questions of the effects of constitutions in practice; claims that would depend on the differing contexts of individual constitutions (Galligan and Versteeg 2013, 11).

Expressions of value need not be more than aspirational statements or window dressing. They might serve as "mission statements" guiding the polity with some holistic purpose as to the nature of its commitments (King 2013, 81). A constitutional text might tell of the goals, ideals, and normative standards its authors wish the polity to be judged by, both by others and by their own future (Murphy 1993, 10). At the very least, constitutions can generate symbolic value which can be utilised to cause a sense of legitimacy and popular support for the state (Ginsburg and Huq 2016, 16).

2.1.2 Enshrining rights

Furthermore, it is often considered a central function of constitutions to lay out the rights (and duties) of the citizens (see for instance Elster 2015, 439). These can be vertical, prohibiting governmental interference in the lives of citizens, or horizontal, though the first type is much more widespread (Elster 2015, 440). In subsequent waves of constitutions-making rights provisions have changed in terms of substance. In addition to the so-called "first-generation rights", which enshrined the rights to freedom of speech, freedom of assembly among others, "second-generation rights" enshrined social, cultural, and economic rights. "Third-generation rights" are considered as vaguer and centres around positive rights such as the right to development (Elster 2015, 440).

2.1.3 Establishing rules of the game and the institutional landscape of the state

Another constitutional function is the determining and regulating of government machinery through the structuring of institutions and of their relations to one another (Elster 2015, 439). This can also be seen as a component in the understanding of constitutions as *coordination devices*. The establishment of institutional power structures, then governing in accordance with the constitution, and the acquiescence of society to the resulting coordination, enables effective governing (Galligan and Versteeg 2013, 23). Constitutions also establish rules for amendment and map out the potential for suspension of the constitution in times of crisis (Elster 2015, 439).

Constitutions can act as stabilisers, by existing as credible commitment instruments, which allows for predictability within systems and therefore economic growth. Constitutions thus provide solutions to problems of information and coordination (Hirschl 2013, 161). This

constitutional function is thought to operate in both democratic and autocratic regimes. Albertus and Menaldo (2013) indeed find that constitutions formed under dictatorships can foster economic growth and increased survival of the autocratic coalition.

Hardin (1999) offers an understanding of constitutions as coordination devices by creating a set of governing institutions, which again constrain certain behaviour. This role of enabling allows for organisation within society, whereas the second role is to simultaneously block certain forms of coordination. Instant coordination ought to be blocked, as the alternative in many cases would be mob rule (Hardin 1999, 82). Hardin seems here to equate the constitutional system, or the constitution, with constitutionalism; coordination might then arguably be seen as an effect of constitutionalism, which is by Hardin understood as enabling and restraining government power in democratic systems. His argument was also a rejection of constitutions as contracts. "A constitution is not a contract; indeed it creates the institution of contracting, which would be *de facto* impossible without a constitutional or other strong order to back it" (Hardin 1999, 87).

2.1.4 Social contracts

The use of contracts as metaphors for understanding constitutions and explaining their legitimacy in their role of binding autonomous individuals have a long history. It has not been without its fair share of critique, as previewed in the previous section. Contracts typically require agreement from all involved parties, a neutral third-party enforcer and are expected to be fulfilled rather than permanently endure (Ginsburg 2013, 182-184), though it might be worth to note that as the median constitutional lifespan is 19 years (Elkins, Ginsburg and Melton 2009, 129) they cannot be considered as permanent institutions. Ginsburg (2013, 183) offers contract theory as a supplementary tool to coordination theories. The argument is that contract theory helps explain how those actually involved with the drafting of the document bargain, and therefore gives a practical insight into the processes of how constitutions are formed as opposed to the theoretical or normative foundation of constitutions.

One such practical aspect is the contents of constitutions. If constitutions operate like contracts, one can expect the drafters to be experts of some kind, representing stakeholders, and drawing on templates of already existing charters and agreements to write their own document

(Ginsburg 2013, 196). And while there may be no ultimate third-party enforcement, there are international forces and agreements exerting pressure on states to include or uphold certain provisions in their constitutions (Ginsburg 2013, 186-187; Alter 2014).

Constitutions thus perform according to contract theory if we understand them as negotiated by a few powerful agents, then acquiesced to by the people at large for reasons either explained by coordination theories or contract theories (Ginsburg 2013, 185). This perspective on constitutions thus links their functions as coordination devices with their functions as elite bargains.

2.1.5 Constitutions understood as elite bargains

Built on an understanding of elites as rational actors seeking to protect and maximize their self-interest, it is posited that political elites will play an instrumental role in the processes of constitution-making and reform (Galligan and Versteeg 2013, 19). "This *strategic-realist* approach is premised on the notion of constitutional law as a form of politics by other means" (Hirschl 2013, 157). This understanding then offers an alternative to those that considers constitutions and constitutional reform to be an ideational victory or a response to the need for coordination.

If one applies such an understanding of constitutional drafting and change, it might at first glance seem counter-intuitive that some less-than-democratic states ends up with some limitations on power baked into their charter. This is assumed explained, however, by how secure elites feel in their position and whether they feel the need to secure themselves for a time after leaving office or other positions of power (Hirschl 2013, 167). If those elites feel rather safe in their continued holding of power, they might seek fewer restraints in the documents.

Brown (2002, 101) gives us another example of how constitutions can perform as bargains between elites; he considers the handling of slavery in the U.S. Constitution to showcase how personal interests of powerful actors and compromise can trump long-term ideological considerations.

“Constitution making also illustrates the variety of motivations that can animate social agents, and notably the interplay among interest, passion, and reason” (Elster 2015, 437). When constitutions perform the function of elite bargaining, they might be considered as more compromise, indecision, and chance, rather than careful deliberation (Galligan and Versteeg 2013, 20). Elster (2015, 447-450) argues that while constitutions and constitutional design matter, they might not matter in the same ways or to the same degree as its designers might have thought.

Certain intended effects can fail to materialise, and unintended consequences can arise. This might lead to the need for renegotiation, or amendment, to borrow from contract theory (Ginsburg 2013, 194). Or such unintended consequences might lead to more constrains on powerholders than the drafters intended for. Brown (2002, 103) gives the example of how European constitutions of the nineteenth century, written to serve short-term elite bargains, had served principles of constitutionalism.

2.1.6 Limitations on government

Often considered by some as the most important function of a constitution, is the limitations put upon government by the general outline of principles and provisions any future government is to abide by. Constitutions thus simultaneously give form to governments while limiting its exercise of powers (Sartori 1962, 856).

One of the central premises this paper rests on is that the existence of a "correct" way of exercising power laid out by the constitution will raise the costs of acting "incorrectly", even if there is a dearth of obstacles in the form of liberal democratic institutions. This cost might be nowhere near high enough to prevent arbitrary exercises of power, or it might in fact be negated by a document allowing these very things. But as there seems to be a convergence in constitutional form across democratic and authoritarian regimes, even if this does not translate into a convergence of functions (Elkins, Ginsburg, and Melton 2014, 162), there are fewer documents without some forms of constraints or limitations on power.

Elkins, Ginsburg, and Melton (2009, 39) argues that while establishing the rules of the game inevitably provides some form of restraint on executive behaviours, they are fundamentally,

albeit subtly, different from constraints stemming from principles of constitutionalism. This distinction is what Sartori (1962) argues for with his adaption of Loewenstein's framework into the categories of *garantiste*, *nominal* and *façade* constitutions. His argument is that the function of constitutionalism was inherent to the original understanding of what a constitution was supposed to do. The document would, along with a corresponding set of institutions, "[…] restrict arbitrary power and ensure 'limited government'" (Sartori 1962, 855). This garantiste understanding differs substantially from the nominal constitution, which was fully complied with, but nonetheless was without any actual restraints on power. A nominal constitution is a power map, indicating where power already exists. The third category of sham, however, contains the provisions of a garantiste constitution but fails to comply with them (Sartori 1962, 861).

This contrasts with other approaches to constitutionalism, such as that of Brown (2002). His view of constitutionalism as a function can emerge from constitutions not at all designed for this feature has been pointed out earlier. Constitutionalism, and it's differing definitions, are the topics for this next section. Particular attention is paid to the discussions centred around the inclusion or exclusion of democracy as a prerequisite of the constitutionalism concept.

2.2 Different understandings of constitutionalism

Constitutionalism is thus one of those concepts where each author seemingly has their own, slightly different, understanding as to the meaning of the word. Some of this assumed discrepancy might simply stem from how expansive a definition is applied and what limitations on government is thought to actually entail. For instance, there are scholars who consider liberal democracy as an integral prerequisite to the functioning of constitutionalism (for instance Hardin 1999), while others exclude it from the concept altogether (for instance Brown 2002).

"[…] in all its successive phases, constitutionalism has one essential quality; it is a legal limitation on government; it is the antithesis of arbitrary rule; its opposite is despotic government, the government of will instead of law" (McIlwain 1947, 21-22). This McIlwain (1947, 22) holds to be the most ancient part of constitutionalism, and the constant, if not the most important part. Real protections against arbitrary and despotic government, he argues, is

those of the ancient legal limitations and what he calls the modern political responsibility (McIlwain 1947, 142). Even definitions as seemingly simple as those concerned with limitations on government generate debates. Tushnet (2013a, 36-39) argues that a narrow reading of the first definition[2] would allow for constitutionalism to exist in regimes decidedly not committed to thicker definitions of the concept. The addition of modern political responsibility to all of the people by those elected, however, changes his assessment of McIlwain's definition to what he calls the creation of a dichotomous conceptualisation of constitutionalism. The alternatives must be either despotism or constitutional democracy (Tushnet 2013a, 36).

> "The element of constraint means that neither anarchy nor a totalizing concentration of power (in one, a few, or many hands) is consistent with constitutionalism. Between these two poles, however, a range of constitutionalist politics or political systems is possible. A constitutionalist system will include three essential elements: (1) institutions authorized by and accountable to the people (both in the regular operation of government and, perhaps, in the making of the constitutional order); (2) some notion of limited government (whether by the designation of purposes for governmental action, the specification of rights, or the allocation of authority among institutions); and (3) the rule of law (i.e., the regularization of processes by which public norms are made and applied)." Brandon 2015, 2

This definition opens for such a pluralization of constitutionalism that Tushnet (2014) argues for, as a fully democratic system of governance is not at all necessary to satisfy these demands. Brown (2002, 9) gives us the historical examples of British, German, and American constitutions; these were well-known for excluding the majority of its citizens from the democratic processes. Many scholars also include rights provisions as an important aspect of limitations placed on the government, though it has also been argued that the presence of rights is of an empirical nature, rather than a definitional necessity (Brown 2002, 9).

McIlwain (1947) observes that the understanding of constitutionalism has changed over time. As we can see from the few definitions discussed in this section, however, at its core, it seems the concept of constitutionalism centres around limitations on power. The most minimalist

[2] Along with the notion that "[a]ll constitutional government is limited government" (McIlwain 1947, 21).

definitions might then settle for just this, whereas the more maximalist definitions embroider on the how and the why. This is not uncontroversial either, though. Holmes (1988, 227) argues that "[i]n general, constitutional rules are enabling, not disabling; and it is therefore unsatisfactory to identify constitutionalism exclusively with limitations on power."

2.2.1 Liberal constitutionalism

Liberal constitutionalism can be defined in manners that at first do not distinguish it from, let's call them, "less liberal" definitions. These definitions might focus on enabling and/or restraining government power through the law. Whether implied or explicitly stated, however, is the sense that constitutionalism is something that can only happen in conjunction with democracy. An example of this is in how Murphy (1993, 9) describes the constitution acting as a guardian of rights, assessing that "[i]nsofar as it is authoritative and embodies constitutionalism, it must protect substantive rights by limiting the power of the people's freely chosen representatives."

Liberal constitutionalism is often associated with ideas of separation of powers, democratic elections, and rights enforceable by law. Individual autonomy, and therefore individual rights, are central to this concept of constitutionalism (Thio 2012, 134). The rule of law is also a concept commonly associated with this type of constitutionalism. "The 'rule of law,' liberally understood, implies more than simply that all actions of the state should be legal" (Slagstad 1988, 106). Laws must, according to this tradition of thought, be a formalization of norms, fulfilling specific criteria such as that of generality. Not just any command or legislative measure can be interpreted as proper laws. What characterises this viewpoint is the idea that the citizens need protection from the state, though it can be argued to presuppose an absolute monopoly of power by the state (Slagstad 1988, 108).

There are, interestingly enough with regard to liberal constitutionalism, tensions inherent in the relationship between democracy and constitutionalism. "For those who believe that there is a conflict between constitutionalism and democracy, the tension stems from the fact that constitutions remove certain topics from public scrutiny and review" (Sunstein 1988, 338). Constitutions then constrain the will of the people. Sunstein (1988) argue there is no *inherent* incompatibility in the relationship, however, different forms of democratic conceptions and

understandings of constitutionalism may not be compatible. The different provisions in the constitution have different functions; "[w]e may distinguish, for example, between structural provisions and rights provisions" (Sunstein 1988, 327). Rights provisions remove subjects from the purview of majoritarian control, but some rights are arguably in service of democracy, such as the right to vote or freedom of speech and can arguably be insulated to the advantage of democracy. Other rights provisions can be seen as a formulation of mistrust in democracy. Structural provisions can potentially mitigate pathologies associated with democracy and improve upon its functioning by insulating politics from factional tyranny or self-interested dealing (Sunstein 1988, 328).

2.2.2 Constitutionalism without liberalism

There are many different adjectives associated with constitutionalism outside of liberal. Some such descriptors might be authoritarian, mixed, monarchical, illiberal, mere rule of law. What they have in common is that unlike anti-constitutional regimes, they do not lack limiting constitutive norms (Thio 2012, 136). "Generic constitutionalism" is a conceptualisation of constitutionalism which has constraining and empowering government through laws and institutions as its main objective (Thio 2012, 134).

Restraints on government can come in different forms, as briefly and implicitly stated in the previous subsection. Both institutional rules hindering self-dealing and rights provisions protecting citizens from state intervention can constrain. The need for stability and credibility related to promises made by government are valuable in non-democratic regimes, just as they are in democratic regimes (Albertus and Menaldo 2013). There is also the possibility of regimes which might actively seek to use to law to deal with perceived challenges in ways that do not seem all that limited, but do follow principles of generality, publicity, prospectivity and respect independent judges or courts. Such regimes do not seem fully compatible with most definitions of constitutionalism but are not completely arbitrary or unrestrained either (Tushnet 2013, 39).

Restraint might also come about as unintended consequences of a constitution-making process. Constitutions, or constitutional provisions, may fail to produce intended effects, or prove to work quite differently from what authors envisioned (Elster 2015, 447-449). Several Arab

regimes have suspended constitutions, which were drafted according to their rulers' preferences, as some provisions gained force in unforeseen manners (Brown 2002, 93). Constitutional arrangements and compromise made to forestall uprisings or revolutions have provided examples through history of documents having a constraining effect where this was not the goal of the drafters, and there are possibilities in of accidental constitutionalism within arrangements of institutional autonomy (Brown 2002, 103, 199-200).

2.3 Conceptualizing constitutionalism

As with any concept with differing interpretations, constitutionalism can be tricky to operationalise. Too wide, or minimalist, of an operationalisation risks landing us near the top of the ladder of abstraction. Such a conceptualisation would be close to useless for the purposes of measuring present here. Given the theoretical differences in approaches to constitutionalism, it is still the goal of this paper to not exclude those more minimalist definitions. This is also important when considering measurement; risking ending up near the bottom of the ladder of abstraction excludes potentially interesting cases from the gathered information. "It might be better to acknowledge the possibility of a plurality of constitutionalisms ranging from an idealized liberal constitutionalism to something short of pure authoritarianism" (Tushnet 2014, 116).

The view adopted here is therefore that the practice of constitutionalism ought to be seen as a matter of degrees. Chen (2014, 2) equates the achievement of constitutionalism to that of rule of law and therein draws on Fuller's "morality of aspiration", which dictates the need to always strive for a higher achievement of an ideal, and thus establishes that constitutionalism is also a matter of degree. This allows for a continuous concept, rather than a dichotomous one. While a dichotomy might be more frequently applied to normative considerations of constitutionalism, as Tushnet (2013a, 36) points out, there are possibilities of other forms of constitutionalism existing somewhere between authoritarianism and liberal constitutionalism. Chen (2014), for example, conceives of genuine constitutionalism, something which can be achieved, and differing types of constitutionalism, such as socialist or communist constitutionalism and hybrid constitutionalism, which are not fully compatible with this first type. These types of classifications allow for a comparison of countries with very differing practices.

A more minimal definition of constitutionalism then, that still encapsulates essential aspects of the concept, would be one of restraining and enabling government power. How constitutions function as a structuring of government power must be subject to an assessment of the distance between the text itself and its de facto environment. "[...] it is relatively easy to make a constitution, but more difficult to put it into practice, to implement it and be governed by it - which is what 'constitutionalism' is about" (Chen 2014, 1). Rather than seeking to measure constitutionalism directly, this thesis attempts to present an argument for starting with the measuring of constitutional compliance.

2.3.1 Compliance as a necessary mechanism of constitutionalism

The central premise to this paper is that compliance is in and of itself a useful concept when considering constitutional functions and the measurement thereof. What functions a constitution might perform, and to what degree, when it is being applied in practice might differ from the functions the same text could have if there was little to no de facto correspondence to the document. Regardless of which definition of constitutionalism is applied, there must be compliance for there to be any practical limitations on power present where the text has stated they ought to be. Otherwise, the constitution is no more than a sham (Tushnet 2014, 11).

Establishing compliance might thus help in distinguishing shams from nominal or garantiste constitutions, even if it does not necessarily differentiate between the latter two. However, the delineation between sham and some form of constitutionalism might not be as clear-cut as mere compliance when taking into account aspirational documents. It can be argued that a constitution is no sham even in cases of great disparity between de jure and de facto as long as the polity’s leaders truly treat the document as aspirational (Tushnet 2013b, 1985).

If one applies a more minimalist definition of constitutionalism, it can be presumed that the presence of compliance is a starting point for the achievement of constitutionalism. It is the mechanism through which we can start seeing degrees of constitutionalism. If a more maximalist approach to constitutionalism is used, the establishment of compliance is still a necessary, if by no means sufficient, condition, and as such measuring it ought therefore still be useful. Considering compliance as a necessary point of departure for constitutionalism

cannot help us distinguish "shamness" from aspirational documents, which must be taken into account when considering the measurement.

The functions of a specific constitution can vary both over time and simultaneously overlap (Ginsburg and Simpser 2014, 8). If constitutionalism is enabling and restraining the government, this then contributes to other functions, such as that of the coordination device. A blueprint for how something operates is only to be followed if it can be assumed other "assemblers" also abide by the same plan. In enabling and restraining government, the constitution is laying out how different groups of society are supposed to coordinate. Compliance then becomes vital also for constitutions understood as coordination devices. Przeworksi (2014) describes it as a functioning as operating manuals, when constitutions describe actual political practice. This function of describing politics as is, is perhaps less limiting than constitutionalism, but will also rely on some weak form of compliance as rules being followed provide actual constraints.

There might be the need for an additional indicator for fully measuring enabling and restraining government power; a measure that captures more than just a match between textual prescriptions and reality. I would argue that a more fully developed compliance measure (the construction of which will be discussed in the next chapter) might capture some of these aspects better and might therefore potentially be a reasonable start at operationalisation of constitutionalism. This is a discussion I will come back to in the concluding and overall remarks for this thesis.

2.4 Previous, quantitative research on constitutions, de facto effects of de jure provisions, and compliance

This section provides a review on some of the quantitative research on constitutions, with particular attention paid to the connection of de jure provisions and their de facto corresponding situations. Some of this literature provides a snapshot into studies on the effect the presence of de jure provisions has on de facto performance. The empirical literature on the effect of de jure rights provisions typically finds little effect of constitutional rights on de facto practices (Elkins, Ginsburg, and Melton 2016, 243). The field is not vast, nor is it as much concerned with compliance as it is with effects. The section detailing these studies will therefore be somewhat

brief. The other research that will be addressed in this section consists of a few articles concerned with the de jure de facto "gap," which are more central to the inquiry of this thesis. The remaining work fall somewhat outside these two main categories.

2.4.1 Effects of de jure rights provisions

The question of whether constitutions are merely parchment barriers, or if they have real restraining effects on government action has been investigated both from a theoretical and philosophical perspective, as seen in the previous sections of this chapter, but also from an empirical viewpoint. Research involved with measuring the effect of constitutional law on state repression has found evidence to suggest some rights respected by the government are associated with the presence of a de jure provision. Davenport (1996) finds that state use of censorship and political restrictions decrease when associated with the presence of freedom of the press and the right to declare state emergency in the constitution. Interaction effects were also found; press freedom and political conflict have less of a chance at seeing repressive measures from the state. The same finding holds for the interaction between states of emergency provisions and political conflict. The study observed 39 countries from 1948 to 1982 (Davenport 1996, 636).

Keith, Tate, and Poe (2009) conducted a time-series cross-section analysis over a 21-year period, covering 154 to 178 countries. Their study claims a decrease in state terror is associated with having provisions for individual freedoms (physical integrity rights), provisions for an independent judiciary and states of emergency provisions in place for 10 years. Specifically, they found public and fair trial provisions to be statistically significant, and that the right to strike was significant in some of the models. No other individual freedom was statistically significant in any way, which was summarised as "somewhat disappointing" and in line with previous research, as Davenport (1996) also had only found three out of 14 provisions to impact state repression. Judicial independence had similar results, where three of the indicators were significant in some of the models. States of emergency provisions, however, were mostly significant (Keith, Tate, and Poe 2009, 654).

Chilton and Versteeg (2014) seek to test whether six different political rights impact government behaviour. Their hypothesis, which they find evidence to support, is that effects

of individual and organisational rights differ from one another. Their theorised explanation for this difference lies in the incentives and means available for organisations to work towards holding the government accountable for constitutional adherence, ensuring survival of the organisations. Such provisions, in contrast to the individual political rights, become self-enforcing. The six analysed rights are the right to form political parties, the right to unionise, the freedom of association, freedom of religion, freedom of movement and freedom of expression (Chilton and Versteeg 2014, 575-576). They find that the right to establish political parties, freedom of religion and to unionise is associated with a statistically positive and significant effect on de facto respect for those rights. Freedom of association had a positive effect, but only became significant in some of the robustness checks. Freedom of movement and expression were not significant. The statistically significant rights also faced a lower probability of severe state restrictions when the corresponding right was present (Chilton and Versteeg 2014, 583).

Ginsburg and Melton (2014) are, unlike the other research in this section, not concerned with rights provisions. The topic of their study is judicial independence. Using observations from the period from 1960 to 2008, covering 192 countries, they find no significant results for the six different de jure judicial independence variables. The focus on their study is between-country variance, as de jure and de facto judicial independence tend to remain relatively stable within-country over time. With covariates removed, some associations became significant, with the authors commenting on a probable correlation between the independent measure and the control variables. When controlling for different levels of regime type, however, they found that authoritarian states with removal and selection procedures that increase judicial autonomy had an effect on the de facto judicial independence compared to other authoritarian states (Ginsburg and Melton 2014, 205). They further find that the effect is stronger if there are additional veto players who can check the executive in the processes of selecting and removing judges. (Ginsburg and Melton 2014, 206).

Metelska-Szaniawska and Lewkowicz (2021) look at the de facto protection of constitutional rules in post-socialist countries in Europe and Asia, and whether de jure provisions have any effects on them. While there is only one significant, unconditional and positive effect related to the freedoms of assembly and association, they do find significantly positive effects on de facto rights protection conditional on political competition, judicial independence and robust civil society. The aggregate measure of de jure rights is also insignificant.

The last study included in this subsection is related to assessing constitutional performance, but with a slightly altered focus. Elkins, Ginsburg, and Melton (2016) analyse the effect of constitutional age on de facto rights protection. De jure rights are accounted for, but the focus of the article is on testing for effects of age. They hypothesise differing effects for different contexts, with maturation, decay, and stasis all possible outcomes. The results indicate a maturation effect for rights provisions, which is at its most pronounced when age surpasses fifty years old, in authoritarian regimes with a relatively high degree of judicial independence, and in states that have performed poorly over time (Elkins, Ginsburg, and Melton 2016, 265).

2.4.2 The de jure-de facto «gap»

"Efforts to measure constitutional compliance on a global scale has yet to materialize" (Law and Versteeg 2013b, 870). This observation seems to have animated the article attempting to assess whether constitutions were shams with regards to rights provisions. Law and Versteeg (2013b) use data about the content of constitutions and actual practice in corresponding situations to measure the distance between text and reality. They do so by assigning a score of 1 or 0 if a particular provision is present, and by assigning 0, 0.5, and 1's to actual practice across the de facto rights protections. The resulting gaps provide a measure of which states overperform, underperform or mostly comply.

This gap is used for the creation of a typology of sham, weak, modest, and strong constitutions (Law and Versteeg 2013b, 883). Underperformers are the sham constitutions, which promise more than they deliver. Weak constitutions promise little and uphold little, whereas modest constitutions overperform by promising little but upholding many rights in practice. At the other end of this typology are strong constitutions that promise much and delivers much. Law and Versteeg's (2013b) article is a massive work and has been very instructive for how the compliance index is formed, even if it does not mirror this gap analysis.

Law and Versteeg (2013b) also run a regression on their measure, seeking to identify causes of constitutional compliance. They find that comprehensiveness of the constitution, a statist constitutional ideology, population size, ethnic fractionalisation, and civil war are all associated with increased violation of de facto rights. Democracy and GDP per capita, on the other hand,

is associated with fewer violations. Western Europe and North America also generally tend to be associated with fewer rights violations than other regions, particularly North Africa and the Middle East perform poorly relative to Western Europe and North America (Law and Versteeg 2013b, 928-930). Characteristics of a country's formal legal structure, interstate war, ratification of human rights treaties, constitutional age, common law tradition, and judicial review all lack statistical significance in their model. The presence of limitation clauses for rights are not proven to have any effect on compliance (Law and Versteeg 2013b, 930-934).

Another work of theirs, is a chapter for a book concerning constitutions in authoritarian regimes (see Ginsburg and Simpser 2013), where they further break down trends in compliance among different strains of authoritarianism. They find that both military and monarchical regimes adopt weak constitutions at a higher rate than civilian authoritarian regimes. Military regimes are also more likely to adopt modest constitutions, compared to civilian regimes (Law and Versteeg 2013a, 185-186). Civilian regimes then tend more towards sham constitutions as pertains to rights provisions. These differences show how measurements at the state-year level for constitutional compliance might give us ideas about causal mechanisms to explore through other means. For instance, why do we see this difference between strains of autocratic regimes? Are civilian regimes hoping for another source of legitimacy from a constitution that has more rights, whereas military and monarchical regimes have their legitimacy from other sources and thus choose to avoid the costs associated with "shamness" (Law and Versteeg 2013a, 186-187)?

Another study done with a similar methodology for mapping the gap between de jure and de facto rights concerns post-socialist regimes and measuring determinants thereof (Metelska-Szaniawska 2021). The findings are a strong path dependence for the size of the gap, and that levels of democracy, presence of political conflict, constitutional age and comprehensiveness of the constitution all affect the gap. Promising more rights leads to greater gaps, as does political conflict and increasing constitutional age, while a higher level of democracy decreases the gap (Metelska-Szaniawska 2021, 188-189).

Elkins, Ginsburg, and Melton (2009) collected the dataset that is used for de jure indicators in this dataset. Their book about constitutional endurance is not the only one resulting from their project, but it also contains very preliminary analysis of cross-sectional data to assess the gap between de jure and de facto rights protection and legislature power. They find that provisions

regarding legislative power see a higher degree of compliance than does right provisions (Elkins, Ginsburg, and Melton 2009, 29-31; 53-55).

3 Building an index – how to measure compliance?

This chapter continues the discussion of conceptualisation, with the goal of operationalisation and measurement. The chapter is structured as follows. Firstly, a discussion of choices made in the conceptualisation for constitutionalism and constitutional compliance. The second part will outline which variables are selected and how they are paired to map the gap between de jure and de facto. The coding scheme applied is also explained. Lastly, the chapter will contain a discussion of validity, potential problems and future improvements of the measurement.

3.1 Conceptualisation

Any conceptualisation needs to be anchored in the empirical analysis of the phenomenon if it is to be anything but an exercise in semantics (Goertz 2012, 4).

The elements conventionally understood to be part of an empirical concept are labels, attributes, indicators, and phenomena (Gerring 2012, 116). Another, closely related, way of conceiving of concepts are in multiple levels: theoretical definitions as a starting point, the constitutive dimensions of the basic concept and lastly the indicators or data (Goertz 2012, 6). The theoretical proposition is, for the purposes of this thesis, that *constitutionalism, as a constitutional function, restrains and enables government power*. This removes the concept from potential attributes, such as liberal or authoritarian, which is desirable when attempting to measure all states against the concept.

Only restraining government is here considered as a subminimal definition, as it fails to recognise enabling as another essential attribute of constitutionalism. One could argue that the addition of this attribute would still result in a subminimal definition, given how potential attributes such as democratic government has been left out. My stance is rather that adding democracy would result in a maximalist definition, as the aim of this conceptualisation and is not to capture the phenomena of liberal constitutionalism. The attributes of constitutionalism will therefore here be considered to be the restraining and enabling of government. What is of importance is not *where* this restraining and enabling happens, in terms of specific provisions or parts of the constitution, but that there is a sufficient quantity of it.

The logic of family resemblance structure is thus applied to this concept of constitutionalism, though I would like to present an argument that whether one would prefer a less minimalist definition of constitutionalism than the one presented here, it is fully possible to restructure the concept and apply additional attributes or constitutive dimensions for necessary and sufficient conditions to be considered met. My argument would be that enabling and restraining government power is no matter the structure an essential part of constitutionalism, and functions through the mechanism of constitutional compliance. Compliance becomes a sort of overarching indicator for constitutionalism. This should remain constant regardless of what other attributes might be desirable to add into the concept.

Compliance with the document is also the mechanism through which the constitution might function as an efficient coordination device or mission statement and is as such not exclusive to constitutionalism but rather its own concept, whose presence is central to measuring constitutionalism and other constitutional functions. What is exclusive to constitutionalism are the attributes of enabling and restraining government power. Or rather the attribute of restraining and enabling, as they are here considered to be two sides of the same coin. Setting out rules of say, a sport or a game, certainly limits the players but it also enables the game to be played at all by specifying what is allowed. As such they happen simultaneously. This might seem very much like an exercise in semantics, as it has little relevance to the structure of this concept, but is worth consideration should one seek to change the structure.

Further breaking this "overarching indicator" down means identifying aspects of the constitution which can be considered as enabling or restraining. As such, provisions detailing the colours of the flag are wholly uninteresting in this context. Those might be best left for a study of the symbolic functions and effects of constitutions. After locating provisions indicating the presence of an enabling or restraining function, comes locating real world measures that identify whether any included article is being abided by. Compliance lies within this relationship of abidance.

Goertz (2012, 30) assigns the task of identifying the negative pole of the basic concept, what occupies the middle ground between these two poles, and whether there is a continuity between them. If the positive is a constitution performing the function of constitutionalism, then the negative is understood here as a constitution performing other, or theoretically possibly no, functions. There is no real restraining or enabling of power through the text. "If constitutions

are effective, the gap between textual aspiration and performance will be small; if the gap is large, the constitution should be deemed ineffectual" (Ginsburg and Huq 2016, 10). A sham constitution, as addressed in the previous chapter, would be an apt descriptor for that negative pole. Furthermore, I think of constitutionalism as a matter of degrees, like others such as Chen (2014). These functions might be performed to a very small degree, but if they still are present then that is still sufficient to classify as a very low degree of constitutionalism. Drawing a clear line between when "shamness" stops and when constitutionalism starts is tricky, but this is where the concept of compliance enters.

Constitutional compliance becomes a concept unto itself, necessitating a definition. I understand it to be the connection of de jure and de facto, and this hopefully resonates with how others would think of it as well as per a criterion of conceptualisation established by Gerring (2012, 117-119). Its' constituent part is the matching of constitutional text and practice, and its' negative pole is non-compliance. This is another matter of degrees, and the operationalisation of the border between the two will be addressed in section 3.2.3. In the conceptualisation of constitutionalism, it is compliance that acts like the border between constitutionalism and "shamness". Just like all decisions inherent in this conceptualisation, the placement of this border could be contested. Perhaps even should. There might be an aspect of the constitution which fulfils a function of restraining and enabling, but if the whole does not qualify as a case of compliance it is disregarded. Compliance is only present if sufficiently many aspects of the higher law is respected in practice. It does, however, not specify *which* parts of the document need to be adhered to. This demarcates the concept of compliance as belonging in the structure of family resemblance (see Goertz 2012, 36).

The choice of applying a family resemblance structure to both concepts is arguably necessary for concept-measure consistency (Goertz 2012, 95-98), as one concept functions as part of the measuring of the other, higher level one. If the structure of constitutionalism is changed to one of necessary and sufficient in order to add more conceptual dimensions, the way compliance is structured might also need to change. The same holds if applying a necessary and sufficient structure while considering enabling and restraining as two different attributes. The matching of de jure and de facto would probably need to be divided among the different attributes, depending on what types of constitutional provisions they measure. The averaging method used for the aggregation of the compliance measure in this thesis would certainly also lead to issues of consistency between measurement and concept and would require some rethinking.

Previous research attempting to explore the gap between de jure and de facto or measure "shamness" have largely centred on compliance with rights provisions (see for instance Law and Versteeg 2013a or b, Metelska-Szaniawska 2021). However, as has been pointed out in the previous chapter, constitutionalism can convincingly be argued to be about constructing as well as constraining. I therefore wish to provide a cursory look into compliance for other aspects of the constitution as well, in an argument for extending measurement to all of the constitution when discussing sham or compliance, its effects and/or determinants. In order to fully measure constitutional compliance inclusion of rights provisions is obviously imperative but does not paint a complete picture by itself.

It has been argued that it is necessary to choose between an internal or external perspective when measuring constitutional success (Ginsburg and Huq 2016, 6-10). The perspective adopted here is an external one. The intention is for a measurement that is more easily comparable across polities. However, there will inevitably be a loss of information, for instance, what the intended functions of a constitution ought to be in the eyes of those who wrote and adopted it.

3.2 Constructing a measurement for constitutional compliance

For the purpose of creating the compliance index, I have used data from the Comparative Constitutions Project (Elkins and Ginsburg 2021) and from the Varieties of Democracy dataset (Coppedge et al 2021), as well as from the CIRIGHTS project (Cingranelli, Filippov and Mark 2021; Cingranelli, Richards and Clay 2014). The subset of data that is utilised is limited in time to the years spanning 1980 to 2020 and limited in space to 175 countries. This is due to practical concerns related to mitigating the impact of missing data. The following section will be divided into three parts, where the first consists of a walkthrough of the variables selected for measuring de jure attributes and those chosen for their de facto counterparts. The second section goes into a discussion on the data available and strategies employed to mitigate problems related to missing data. The final part in this section of the chapter is an outline of the logic of the coding scheme applied to create the index.

3.2.1 Selection of indicators

The data for de jure provisions are drawn from the Comparative Constitutions Project's (CCP) third version of "Characteristics of National Constitutions" (Elkins and Ginsburg 2021). The dataset consists of single, multiple, and open-ended response questions detailing characteristics about written constitutions. The dataset also includes identifying information on constitutional systems and events. The data for the de facto descriptors are taken from the eleventh version of Varieties of Democracy's (V-Dem) dataset (Coppedge et al 2021). The sole exception is data for disappearances and extrajudicial killings for a de facto measure of violations of prohibitions of arbitrary arrests. This data is gotten from the CIRIGHTS Data Project's dataset (Cingranelli, Filippov, and Mark 2021; Cingranelli, Richards, and Clay 2014). The compliance measure is structured in four components and a total of twenty-four subcomponents divided across these components. How many subcomponents each component has is largely driven by data availability and consistency in the matching of de jure and de facto indicators.

Table 3.1: An overview of selected de jure provisions and the indicators chosen as de facto correspondent, with example of coding procedures

De jure indicator	**Corresponding de facto indicator**
For component regarding the executive	For component regarding the executive
Does the executive have the power to initiate legislation? 0 = No 1 = Yes 0.5 = Not specified or mentioned in text For all indicators of executive power, both head of state and head of government are accounted for it the system has more than one executive. If one is in compliance, while the other is not, a score of 0,5 is awarded instead of 0 or 1.	Does the executive propose legislation in practice? 0 = Yes 1 = Yes, but this power is shared 2 = No Compliance coding: De jure is 1 and de facto is 0 or 1 = 1, De jure is 1 and de facto is 2 = 0, De jure is 0 and de facto is 0 or 1 = 0, De jure is 0 and de facto is 2 = 0,
Does the executive hold veto power?	Would the executive be likely to succeed if they took actions to veto legislation?

Can the legislature override the veto, and if so, with what percentage of the vote?	
Does the executive hold the power to dismiss the legislature?	Can the executive dissolve the legislature in practice?
Does the executive hold the power to dismiss ministers or the cabinet?	Can the executive dismiss ministers in practice?
Total number of subcomponents:	4
For component regarding the legislature	For component regarding the legislature
Do the legislature hold the power to investigate the executive?	Do the legislature investigate the executive in practice?
Do the legislature hold the power to question members of the executive, or does the executive have to report regularly to the legislature?	Do the legislature regularly question the executive branch in practice?
Total number of subcomponents:	2
For component regarding the judiciary and public administration	For component regarding the judiciary and public administration
Are the central judicial organs declared to be independent?	Is the high court considered independent in practice?
Does the constitution provide for an electoral commission to oversee the election process? (electoral courts not considered as EMBs for this indicator)	Does the electoral management body have autonomy from the government?
Does the constitution provide for meritocratic recruitment of, or non-discrimination in the hiring of, civil servants?	Are appointment decisions in the state administration based on personal and political connection, as opposed to skills and merit?
Does the constitution provide for an ombudsman, an attorney general, or other related oversight bodies?	If the executive branch engages in unconstitutional, illegal, or unethical activity, will an oversight body, other than the legislature, investigate and issue an unfavourable report or ruling?
Total number of subcomponents:	4

For component regarding rights	For component regarding rights
Does the constitution freedom of assembly?	Is there a right to peaceful assembly?
Does the constitution provide for freedom of association?	Is there CSO oppression by the state?
Does the constitution provide equality for, or prohibit discrimination based on gender?	Is there exclusion by gender?
Does the constitution provide for a right to health care, or a state duty to provide health care?	Is basic health care guaranteed to all?
Does the constitution prohibit arbitrary detention or unjustified restraint?	Were there disappearances? Were there extrajudicial killings?[3] *Data from the CIRIGHTS Data Project*
Does the constitution provide for access to justice? Answered "Yes" if 4 out of these 9 provisions are present: Right to a defence, or counsel, right to a fair trial, right to appeal, public trials generally required, presumption of innocence, right to an interpreter or language that accused can understand, right to speedy trial, right to redress, or prohibition of punishment by laws enacted ex post facto	Is there access to justice for men? Is there access to justice for women?
Does the constitution provide for freedom of movement?	Is there freedom of foreign travel and emigration? Is there freedom of domestic movement for men? Is there freedom of domestic movement for women?

[3] Measuring unjustified detention by looking at disappearances and extrajudicial killings arguably only accounts for extreme cases of unjustified restraints, but these measures have also been used in other quantitative research regarding compliance with constitutional rights, such as Law and Versteeg (2013a).

Does the constitution provide for the right to form political parties?	Are there barriers to forming political parties?
Does the constitution provide for freedom of the press?	Is there government censorship of print or broadcast media? Is there harassment of journalists?
Does the constitution provide for a right to own property?	Do men enjoy the right to private property? Do women enjoy the right to private property?
Does the constitution provide for freedom of religion?	Is there freedom of religion?
Does the constitution prohibit slavery or forced labour?	Are adult men free from servitude and other kinds of forced labour? Are adult women free from servitude and other kinds of forced labour?
Does the constitution provide for freedom of speech or expression?	Is there freedom of discussion for men? Is there freedom of discussion for women? Is there freedom of academic and cultural expression related to political issues?
Does the constitution prohibit torture?	Is there freedom from torture?
Total number of subcomponents:	14
Total number of provisions measured:	24

Earlier research on the de jure de facto gap has primarily been centred around rights provisions (see for instance Law and Versteeg 2013a or b; Metelska-Szaniawska 2021). The inclusion of different rights in such studies have largely been decided by data availability. This is also largely applicable here. Choices for de jure and de facto variables have been made by many readthroughs of the respective codebooks. The overarching theoretical concern for variable choices has been the conceptualisation of compliance understood as a component of enabling and restraining government power. Restraining and enabling can certainly be seen in the inclusion or exclusion of different rights provisions, in the specified powers of different governmental actors, and relatedly, also in the interinstitutional checks on power. Provisions concerning national symbols such as flags, anthems and the like have been considered as fulfilling other functions of the constitution than constitutionalism, and therefore disregarded.

3.2.2 Impacts of data availability

There are several variables I would have wished to include in the compliance index, but that I either did not manage to find or where the questions asked and answered for the de jure and de facto measurement were incompatible. An example of the latter would be for a subcomponent concerned with the legislature's ability to remove the executive. The de facto measurement explicitly tells its expert survey to not include instances of removal through impeachment, while the de jure measurements are concerned with whether there are constitutional provisions for dismissing head of state or head of government, under what condition this can happen, and upon whose proposal. A quick look through the data showed that France was coded as allowing for dismissal of the head of state by the constitution, but not in practice. The measures are therefore incompatible as long as the one matching the scores together does not have information on impeachment-rules in every country for the entire time period.

A similar problem arose for the election of head of state and head of government. The de jure variables only indicated whether the executive had inherited the position or was appointed through royal selection, elected by the citizens, or elected by an elite group. For head of government appointment was also an option. Comparing this to de facto indicators proved difficult for any country-year the executive was elected by an elite group. Arguably, the party in a one-party state and the legislature are both elite groups. And how would the military fit into this? Is a royal council an elite group or would selection by that institution qualify as royal selection? No matter which potential elite group selects in practice, how does one assure that it is the elite group specified by the constitution?

Another subcomponent that got rejected was the power of the legislature to initiate legislation. The de jure variable asked for who could initiate legislation with multiple options available for the coders, whereas the de facto variable asked whether approval of the legislature was required for legislation to pass. Being given the power to initiate legislation does not equate to being required to approve for any legislation to pass, even if they are likely highly related. The inclusion of the right to education, as a socioeconomic right, would also have been ideal. The de jure variable did not include any such formulation in variables relating to education.

Whether the constitution contains provisions concerning education does not equal to a guaranteed right.

Other times it proved difficult to locate relevant de facto indicators altogether. The inclusion of de jure term limits and de facto adherence to those would have been preferable, as this must be considered a measure of restraining the power of the executive. Evidence does, however, suggest that while attempts at avoiding term limits are very common, those executives who overstay their time in office often overstay through constitutional amendment or the writing of an entirely new constitution (Ginsburg, Melton and Elkins 2011). Since the year 2000, while a third of all incumbent presidents who reached the end of their term attempted to overstay, none did so by ignoring the constitution (Versteeg et al 2020). The absence of this variable then suggests the compliance scores might be lower, at least for presidential systems, than it would with an inclusion.

Selection and removal procedures for judges is another aspect I would have liked to include in the measurement. These are central conceptions to judicial independence (xxx). Compliance with the indicator for judicial independence included in the measurement is not expected to be widespread. Melton and Ginsburg (2014) find evidence suggesting a higher degree of judicial independence where there is competition between executive and legislative powers in the appointment and removal processes of judges, but no such correlation for constitution containing merely statements of judicial independence.

Other indicators that were excluded include the right to unionise (due to a very high degree of missing de facto data), the power to interpret the constitution, the protection of judicial salaries, whether courts can challenge the constitutionality of laws, whether the public can propose legislation, whether the public can challenge the constitutionality of laws, and rules for constitutional amendment. Procedures surrounding emergency powers and the existence of federal structures could also be of interest in a measure such as this.

Another aspect of limitations to the data availability is the missing data in the datasets. There is data missing from the CCP dataset. Generally, this is for constitutional systems they have not yet coded, which means missingness across all indicators for those country-years and accounts for approximately 8.6 % or 582 of all observations. Some countries were missing for

their entire time-period, and those have been excluded from the index[4]. The CCP dataset already includes imputed data where there have been amendments that have not yet been reconciled by coders or documents have not been located, but not replacements of the constitutional system. This means that most of the missing data are due to changes to the constitutional system, such as replacements, not yet having been coded or a continued effort to reconcile coding decisions already made. Some of the subcomponents in the index are missing at higher rates, typically between 8.6 and 11 %, due to missingness in the various datasets. The subcomponent measuring compliance with prohibitions of arbitrary detentions, however, are missing for 19.15 % of all country years.

3.2.3 Coding the distance between de jure and de facto measurements

The coding scheme applied is as follows:

If a constitution contains a provision, and the state abides by this provision in practice, then the compliance-subcomponent is coded as 1.

If a constitution contains a provision, but the state does not abide by it in practice, the compliance-subcomponent is coded as 0.

If the constitution has the provision, but only abides by it to a certain degree where this is qualitatively different from non-compliance, the compliance-subcomponent is coded as 0.5.

If the constitution lacks the provision, or explicitly leaves it up to non-constitutional law the compliance-subcomponent is coded as 0.5.

To illustrate the logic of this coding scheme follows an example of four different constitutions for one country-year, given (for the sake of simplicity) the total number of provisions measured is 20:

Table 3.2: Illustration of coding scheme

[4] This applies to the countries of Yemen Arab Republic, Israel, Czechoslovakia, and South Sudan.

	Constitution 1	Constitution 2	Constitution 3	Constitution 4
Provisions in constitution	10	15	20	5
How many: Follows (1 point)	8	8	10	4
How many: Does not follow (0 points)	2	7	10	1
How many: Not included (0,5 points)	10	5	0	15
Points	13/20	10,5/20	10/20	11,5/20
Score	0,65	0,525	0,5	0,575

This shows that the index awards a political system with a constitution which contains fewer provisions, but that manages to uphold a bigger portion of those written provisions, a better score than one that has a more specific document but fails to live up to the same percentage of provisions given. The state that performs the worst in the illustration is the one with a constitution containing all provisions measured for compliance (Constitution 3). It is a state that upholds more provisions in practice than all the others, but it also violates more of its own provisions than any other state.

The reason omitted provisions are coded as 0.5, rather than 0, is that these instances are qualitatively different from non-compliance. Non-specificity does not equal a sham document. It ought to still be factored into the score, as more provisions would equate to more restraints or enabling mechanisms. Including the scores of 0.5 for non-present provisions award a lower score for non-specific documents than if the constitution had included more provisions (that were followed by the government). Applied to the opposite end of the scale, this would mean that a country violating a constitution containing many provisions would score worse than a country violating a constitution containing fewer provisions. The logic is that violating many provisions is worse in terms of complying, than violating some provisions and not otherwise being restrained or enabled. Consider Table 3.2, Constitution 1 and Constitution 4 would get

the same score of 0.8 if the score was calculated as provisions complied with divided by number of provisions present. They would score 0.4 and 0.2, respectively, if the score was calculated as provisions complied with divided by total number of provisions included in the index. The former coding strategy would equate Constitution 1 and 4 with one another, and the latter would consider them non-compliant.

For each country-year the total score is calculated by adding together the compliance scores for each constitutional provision and averaging it. The result is a score between 0 and 1, allowing for degrees of compliance. It still enables any future use of the variable as a dichotomous one, as long as a cut-off point is set. It might also be worth noting, when considering the results presented in the next chapter, that the decision to code missing provisions as 0.5 creates a more centred index, than would be resulting from excluding those provisions a constitution does not contain.

There is a theoretical possibility of an "empty" document, should none of the provisions measured in this index be present in one, or more, country-years[5]. Given the inclusion of 24 different provisions, a score of 0.5208 would be the score achieved if a constitution were to only have one of the provisions measured which was then abided by[6]. Partial compliance with one provision, and the rest deemed missing would give the same score as a hypothetically "empty" document. This reflects the attitude that a constitution only partially abided by does not equal compliance. 0.5208 is then the cut-off point indicating the lowest possible degree of constitutional compliance based on these data. For those country-years where there were missing data for one or two de facto indicators this score should be some percentage points, or rather decimal points, higher.

Awarding a score of 0.5 where a provision is somewhat followed is necessarily more difficult in practice than in theory. In some years, certain states will come close to following a provision, whereas some will come closer to failing. Both of these cases would still, arguably, fall under the same category in this coding scheme. There is therefore a loss of information in the

[5] This occurs for Bhutan for the years 1980 – 2004. Canada comes close by registering at two provisions present for 1980 and 1981. While Djibouti and Ghana both have three provisions for a period of years, both New Zealand and Australia are assessed to only have four of the twenty-four provisions for the entire timeseries.

[6] This is calculated by ((23*0.5) +1)/24.

simplicity of the coding choices. In order to make the choice of assigning this value as standardised as possible certain measures have been adopted in the subcomponent-specific coding schemes.

For instance, most of the de facto variables used came from the same dataset (V-Dem) with a similar logic behind coding categories and corresponding values. A specific country-year were for many of the de facto variables assigned values between 0 and 4. For the compliance index, the choice was then made to consider a score of 4 to be in compliance if the provision was present. If the score was 3, then the value assigned in terms of compliance was 0.5. A value of 2 and lower corresponds to 0 in the compliance-subcomponents. Where compliance with a provision is measured by combining several de facto measures, for instance when there were separate indicators for men and women, a score of 1 was accorded for values of either 4 and 4, or 3 and 4. Values of 3 and 3 resulted in 0.5, and combinations involving lower scores were assigned a 0. A more detailed description of the coding scheme can be found in the Appendix. It is important to note, however, that there will always be some arbitrariness related to the assignment of quantitative values for complex phenomena. Others would no doubt assign thresholds for compliance and partial compliance in other ways than I have, and many different approaches may be legitimate for different reasons. In order to be clear on what the index actually measures transparency in terms of coding decisions is sought.

3.3 A discussion on validity, and on consequences of the choices outlined in this chapter

Any attempt at large N-studies within the subject of constitutions and constitutionalism will inevitably be faced with criticism and potentially with charges of meaninglessness of the variables generated or analysed, due to the lack of contextual sensitivity. Within the legal research fields such as comparative constitutionalism and comparative constitutional law there are tensions between those who see constitutions as wholly unique and particular to its own historical, cultural, legal context and those who consider constitutions to consist of universal elements one can draw generalisations from (Hirschl 2014, 197). This split somewhat mirrors the one between relativism and positivism in the social sciences. The approach taken here is that any quantitative large-N project such as this might generate the most insight when seen as a complement to more in-depth qualitative studies. While there are many drawbacks to large-

N approaches, they are still useful in combination with other approaches […] provided they acknowledge their embedded neglect of context and nuances" (Hirschl 2014, 276).

3.3.1 Validity and reliability of the measurement

There are certainly issues with the resulting index. Attempting to identify whether parts of text and reality match up will inevitably overlook differences in constitutional interpretation by courts, the impact of similar or related provisions in different polities, the different routes towards implementation, aspirations and so on. Tushnet (2014, 7-8) gives one such example of where interpretation gives important additional information, by referring to the U.S Constitution's right to bear arms; other constitutions have been interpreted to create a right to bear arms even if such a provision does not exist in the texts. The U.S Constitution is then less of an outlier in practice than what the data sans interpretation might suggest. Other interpretations might bring practice into closer compliance with de jure provisions. The role of international and transnational laws and rulings by the adhering courts is also left behind on the cutting-room floor. As of now, this index lacks the information needed for a more accurate measurement. One advantage of the lack of information about interpretation is that the measure is designed in a way capable of picking up growing divergence from original text with the passing of time, which is a phenomenon of interest in itself (see for instance Strauss 1996 on a discussion of the effects of interpretations on divergence from text). To truly capture this evolution, however, more years ought probably to be included in the scope of the index.

Certain coding decision will also affect the accuracy of the measure. Some of these will be discussed here. For instance, constitutions often contain limitation clauses, in fact Law and Versteeg (2013a, 933) find that they are universal. The argument could be made that the broadest types of limitation clauses, which might state that such and such right shall be restricted by law, might render any limitation in practice as perhaps in compliance with the constitution provided limitations were put forth in law. If so, there would be a lack of information regarding compliance within the data, and the provisions ought to be coded as not present. Similar, but more restrictive limitation clauses might specify for which, though rather broad, conditions or aims certain rights might be curtailed. Such conditions could be "public order," "public morals," or "national security." The same argument as for unrestricted limitation clauses could be applied here as well. However, considering a large number of

provisions[7] as not present, when they are in fact present, because there are limitations on their applicability seems ill advised. Especially so, since they are not found to be correlated with the respect of these rights (Law and Versteeg 2013a, 933-934). The decision was made to ignore the limitation clauses and code these provisions as the rest.

A source of contention in the coding decisions for the index might be in the inclusion of provisions from the preamble. This relates to rights provisions. Preambles are typically a communication of how the authors of the constitution conceive of the project, the past, and the future. They are therefore not always considered as enforceable parts of the constitution. They can be, however (Tushnet 2014, 27), and as such I have coded provisions as present if they were coded as such in the CCP data and their articles were listed as the preamble. This might have led to an inclusion of provisions in the de jure data for cases where there is no sense of such provisions as binding.

Another weakness of the research design is related to its quantitative nature. Measuring the contents of constitutions is very much an exercise in semantics and precise language. There is bound to be some disagreements with quantitative measures of constitutional characteristics regardless, but especially relevant here are those provisions that were labelled as "Other" by the CCP coders. There is uncertainty in the reliability of the coding introduced wherever a provision is coded as such. For those country-years I have read the comments left by the coders in order to categorise a provision as present or not, so as to facilitate comparison with de facto measurements. This is the least reproducible aspect of the research design, and as the decisions have not been reviewed by others, they are all the more questionable for it. Where possible I have consulted the constitutions for those country-years to reduce mistakes made due to lack of textual context, but sometimes documents have proven difficult to access. Whereas most constitutions in force at the moment can be found through the constitute project (Constitute 2021), previous version are not always so easily available. In my anecdotal experience, I found this to be especially true for different amendments rather than for constitutions that were completely replaced.

[7] Law and Versteeg (2013b, 933) find that these limitation clauses are present in every single constitution for at least one rights provision in the year 2006. 43.3 % include a blanket clause of some kind.

An example of such "Other"-related decisions is the question of when freedom of the press can actually be considered as protected by the constitution. This decision impacts coding choices for the country-years in question. For most of those country-years that were coded "Other" for this provision, there were protections of different forms of communication. Whether something along the lines of "all means of communications of ideas, opinions and information is to be respected" can be considered as a guarantee applicable to the press is debatable. I have concluded that it does constitute a provision for freedom of the press for the purposes of this project, but if other were to replicate the index, they might decide otherwise. Arbitrary arrest is another example. Protection from unlawful detention, or right to habeas corpus, have been interpreted to also include those "Other" cases where arrests must be made in accordance with the law. While there has been concerted effort to treat like cases alike, at least within provisions, there is a reason these cases were coded as "Other" in the first place. They are difficult to classify. Luckily, for the sake of both the reliability and validity of the measure, they tend to make up rather few cases of the total[8].

Furthermore, it is up to the individual coders of the CCP project to add comments in order to clarify or expand upon the provisions they are coding. What they find relevant to include might conceivably vary, thus impacting what considerations I have made in the coding of cases. There are also instances, though thankfully rare, where I have found myself in disagreement with the coding decisions made. An example of that would be Mexico's constitution, where the right to form political parties is coded as "Other" for the entire time period of the index. It is made with reference to article 54.IV, and the comment gives no reason to presume a guarantee of this right. In fact, it says it is left to non-constitutional law. In the document amended through 2015, however, article 54.IV deals with election systems for the first chamber. Article 41.I, on the other hand, implies this guarantee when it says only citizens can form a political party. This might be a case of erroneous coding, I might be misinterpreting something, or I might be looking at the wrong version of the document. The Mexican constitution have, after all, been amended 76 times in its life span of 104 years (Constitute 2021). In these instances, where I find myself in disagreement with the CCP score, I have defaulted to their assigned scores and

[8] The number of cases coded as "Other" for freedom of religion was 192 cases, among 6743 total observations. For property rights this number was 204. For freedom of assembly, it was 35. For access to justice, I only manually decided the cases in which fewer than 4 rights were present, but some rights were coded as "Other". This amounted to 862 cases.

comments. This is because there are two coders working individually for each document in the CCP dataset, with a third person reconciling if there are coder disagreements, whereas I am working alone. Such instances do beg the question of whether my coding interpretations for other country-years coded ambiguously are reliable.

There is also a question to be raised regarding the measurement's capture of the concept. Due to issues related to data availability, there are several indicators mapping the adherence of de facto situations to de jure texts that were not included. To fully measure compliance in a manner providing conceptual validity (see Gerring 2012, 161-163) further indicators would need to be added. The further analysis of the index in the following chapters are therefore subject to this (rather large) caveat. This is a limited measure of compliance, which would require substantial improvement in terms of additional indicators and data before any resulting analysis can be validly claiming inferences based on constitutional compliance.

3.3.2 Possible pitfalls and potentials of the measurement, as well as prospective improvements

The purpose of this index, given its numerous imperfections, would first and foremost be to continue a discussion on *how* to gain more generalizable insights into constitutions across different political systems and time, rather than to provide any real answer to such debates. It is also not primarily concerned with normativity; measuring constitutional compliance is a normative endeavour only insofar that it ranks systems from better to worse in terms of more compliance to less compliance. It says little about what a constitution ought to achieve. For instance, the Albanian constitution of 1976 explicitly prohibits ownership of private property. Many would perhaps see this as an unwanted feature in their constitutional system, yet Albania scores 1 for the subcomponent measuring provisions regarding "Freedom of property" during the period this constitution was in place because ownership of private property was virtually non-existent according to de facto data. The exclusion of any normative assessment of what constitutional provisions should do can be addressed by changing the structure of the concepts and adding new dimensions, as mentioned in the chapter's first section. This depends on the purpose a compliance index is thought to have in any research project. There are advantages to being able to measure purely if practice is what the text dictates it should be, such as (mostly)

excluding a function of window dressing without excluding constitutional systems for other reasons.

The index is not designed to account for overperformance of constitutional rights, unlike some previous studies on the de jure-de facto gap in terms of rights provisions (see for instance Law and Versteeg 2013a or b; Metelska-Szaniawska 2021). Constitutional overperformance, often found within states with older constitutions (Law and Versteeg 2013a, 902), within human rights is an important and interesting empirical phenomenon that this measure does not provide further information on, or analysis of. The reason for not accounting for overperformance is the conceptualisation of compliance applied. There is no room for anything that is not included in the text; ignoring the danger of repeating myself, compliance lies within the relationship between de jure and de facto. One should perhaps be careful to say that the constitution, or any single provision, can cause compliance, but constitutional overperformance suggest a causal mechanism outside of what is actually written in the text.

Measuring the "gap" between de jure and de facto constitutions is a complex endeavour. Translations are famously a source of miscommunication and loss of information; comparisons and interpretations might yield slightly differing results depending on which languages the work is done in, and different translations may alter the results as well (Tushnet 2014, 6). Furthermore, what the compliance-score might tell us will depend on other contextual factors, such as the time provisions have been in place or whether they are intended to be transformative or preservative (Ginsburg and Huq 2016, 10-12). Any measure of constitutional compliance is limited in the amount of information it provides us with on its own. As such, the stance adopted here is that a further developed measurement of compliance becomes useful only when combined with other sources of information, preferably of both qualitative and quantitative nature.

4 The compliance index – a descriptive analysis

This chapter presents a cursory introduction to the data of the compliance index, through graphs and tables. The first section gives an overview of some of the variables produced for the index, before the second section delves into some bivariate relationships of the compliance index. The variables introduced in this first section are the compliance index, total number of provisions present in any given country-year, as well as components and subcomponents of the index. These components and subcomponents are as introduced in the previous chapter (Table 3.1). These relationships concern the passage of time within the measurement, the number of articles present, regional differences, and differences by regime type.

4.1 Descriptions of the data

The first variable to be presented is the compliance index. The overview contains some descriptive statistics regarding distribution of scores before this is illustrated in a figure. There is also a figure illustrating the bimodal distribution of scores should the provisions excluded, or perhaps more accurately non-included, in a constitution any given country-year be coded as missing rather than 0.5.

Table 4.1 Descriptive statistics for the compliance index

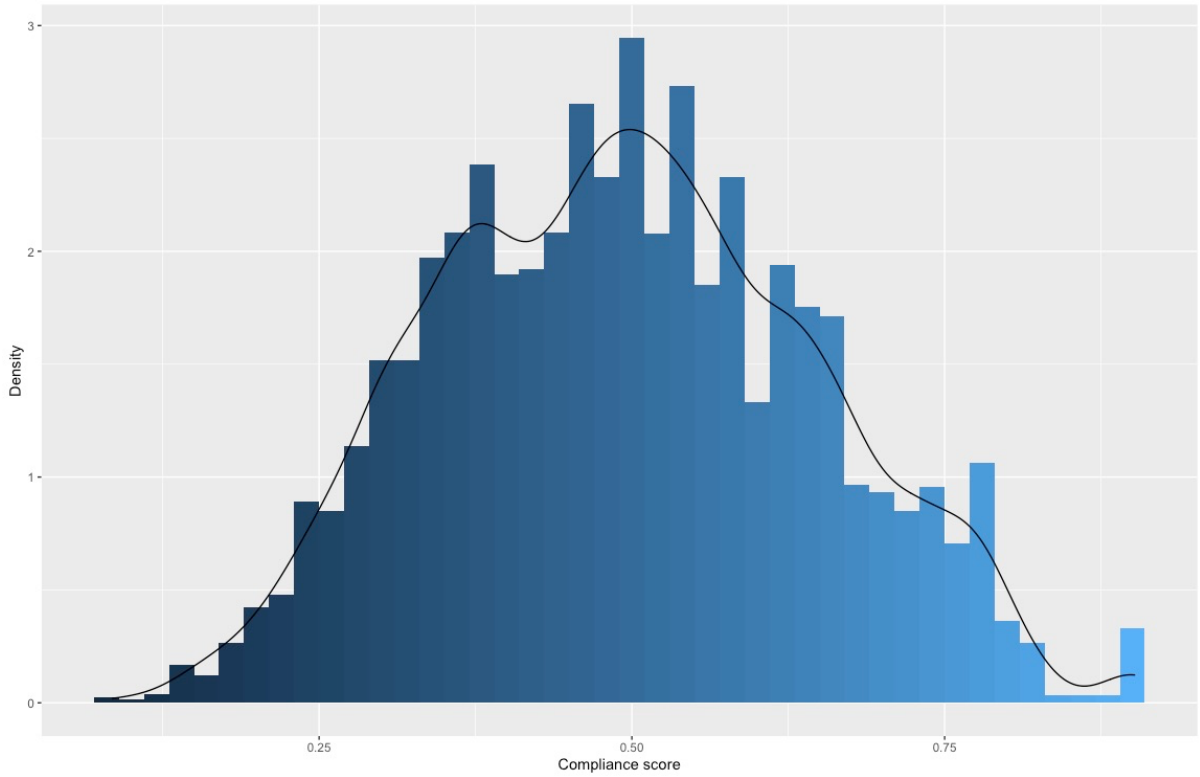
	Compliance
Mean	0.49
Standard deviation	0.15
Minimum score	0.08
Median score	0.49
Maximum score	0.90
Number of valid observations	6161
Skewness	0.14
Standard error skewness	0.03
Kurtosis	-0.48

Table 4.1 presents a descriptive overview of the compliance index. Both the mean and the median fall slightly below what is considered compliant, with both values at 0.49. We can see

that there is a slight positive, and significant, skewness. The kurtosis indicates a slightly heavier tailed result than the normal distribution. This is confirmed looking at Figure 4.1, which shows the distribution of compliance scores for all the years measured for the index.

As indicated in the previous chapter, choosing to code non-existing provisions as 0.5 gives a more centred measure, though with somewhat heavy tails according to the kurtosis-measure. Choosing a different coding scheme for missing provisions, for instance coding de facto situations irrespective of the presence of a constitutional article, would give us a bimodal distribution[9]. For an indication of what this would look like is the total frequency for different values for compliance when all missing-scores are subtracted from the subcomponent-scores (see Figure 4.2). Scores of 0.5, given whenever a provision was indicated to not be present in the constitution for any given country-year, is what is meant by missing-scores. There are more compliant observations than non-compliant, though not by a large margin, while the fewest observations fall in the liminal "almost"-category.

Figure 4.1 Compliance across all country-years



[9] See chapter 3 for a discussion of why this approach was not followed.

Figure 4.2 Frequency-distribution of scores from all subcomponents, subtracted by scores for missing provisions

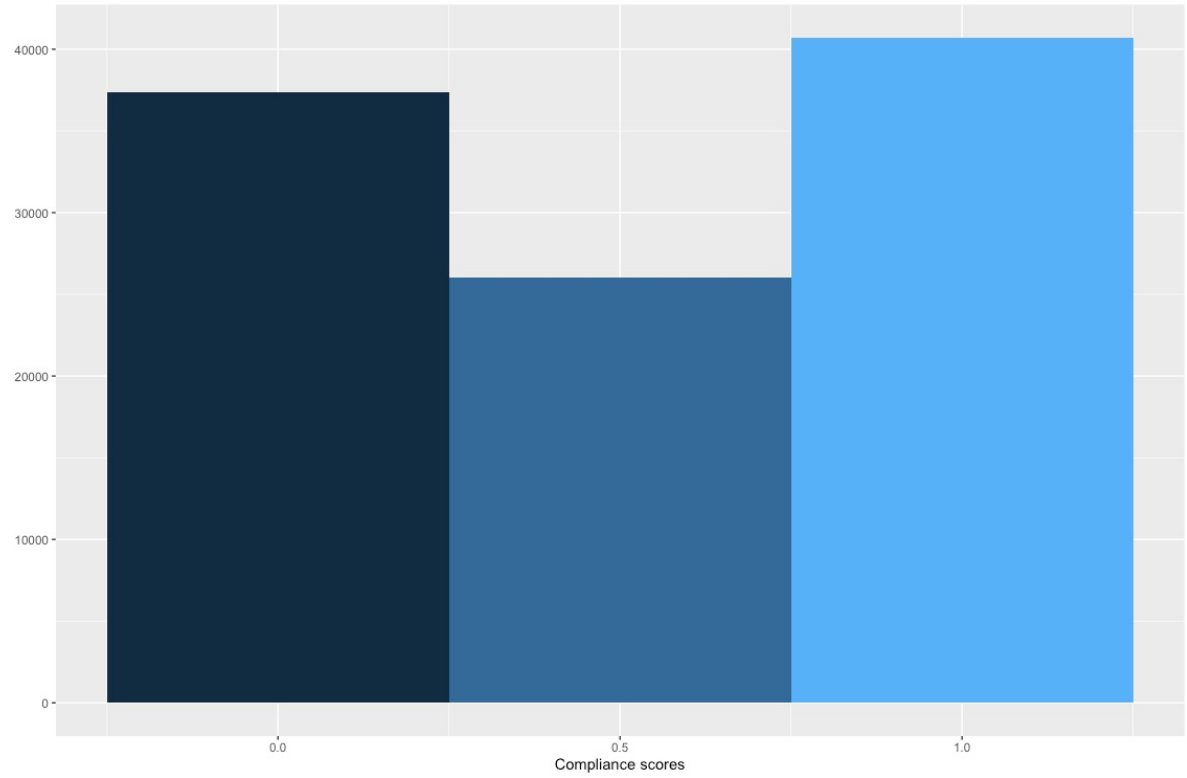


Figure 4.3 provides a look at the different components of the index. It is worth to note the effect of the number of subcomponents measured for each of the four, as the category for provisions regarding rights (bottom-right) have many more columns and therefore much fewer observations for each value. This is especially true compared to the category containing articles concerning the legislature (bottom-left), where there are only five specific values possible. In terms of score-distribution, it is evidently more country-years in which there are compliance with provisions regarding the executive (top-left) compared to respect for provisions regarding the judiciary and public administration (top-right). The former skews towards values indicating compliance, whereas the latter skews heavily towards non-compliance. These distributions include observations for country-years where there is no relevant article.

Figure 4.4 provides another look at the components; this by a boxplot of the distribution of countries within the different components for the year 2020. There is quite the dispersion, as can also be seen in Figure 4.3, with the median-scoring countries ending up above the cut-off for both the executive-component and the rights-component. The legislative-component and

judicial-and-administration-component, however, both have their median below the cut-off indicating some degree of compliance.

Figure 4.3 Compliance for all country-years, grouped by index-components

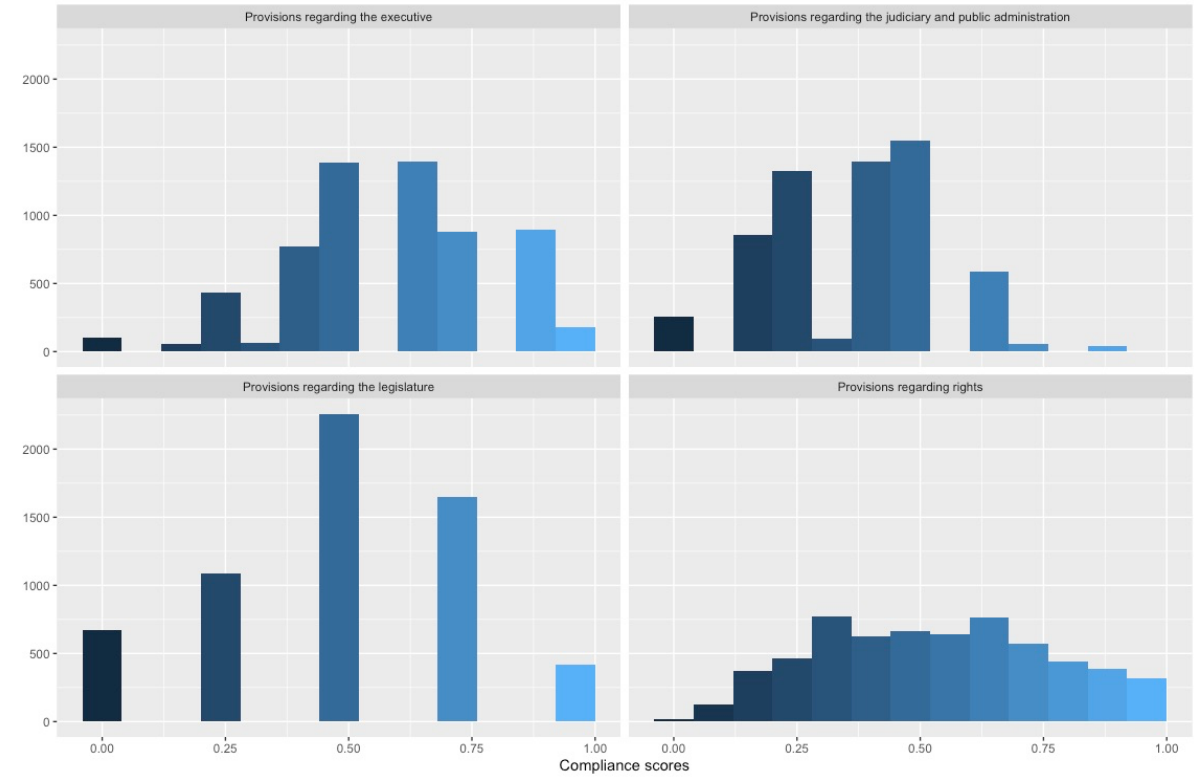
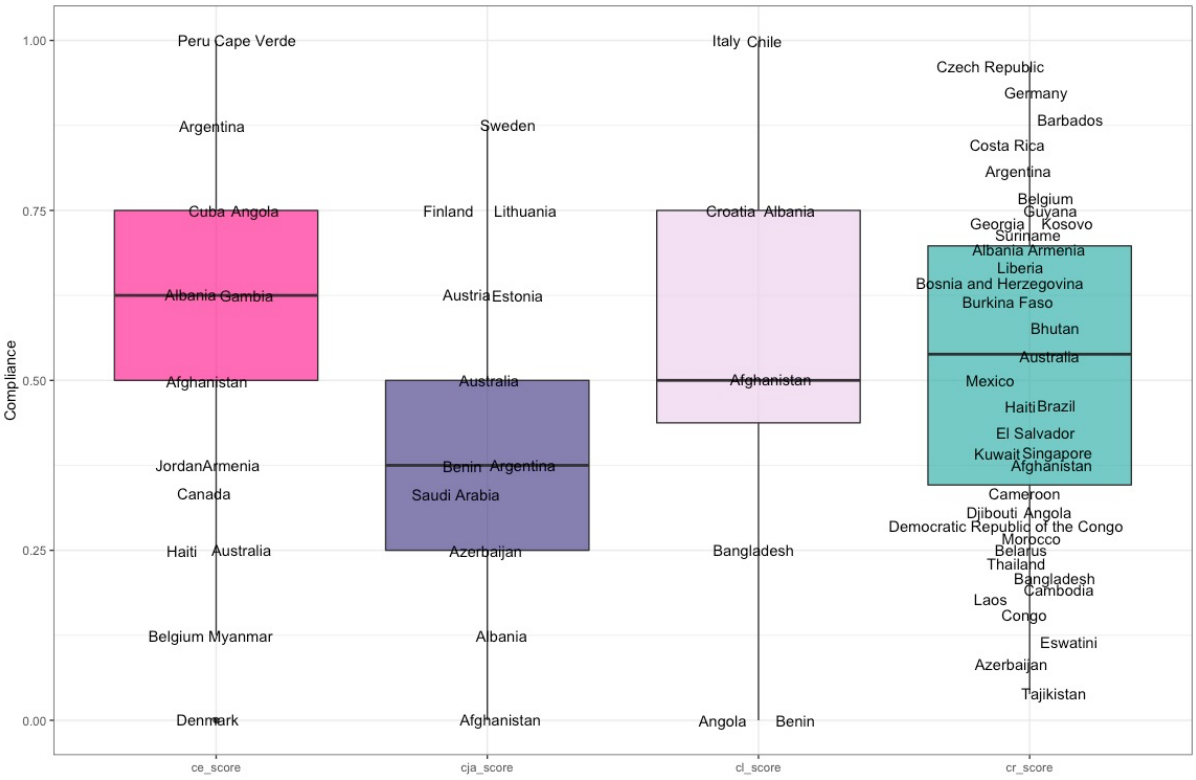


Figure 4.4 Compliance by components, in 2020

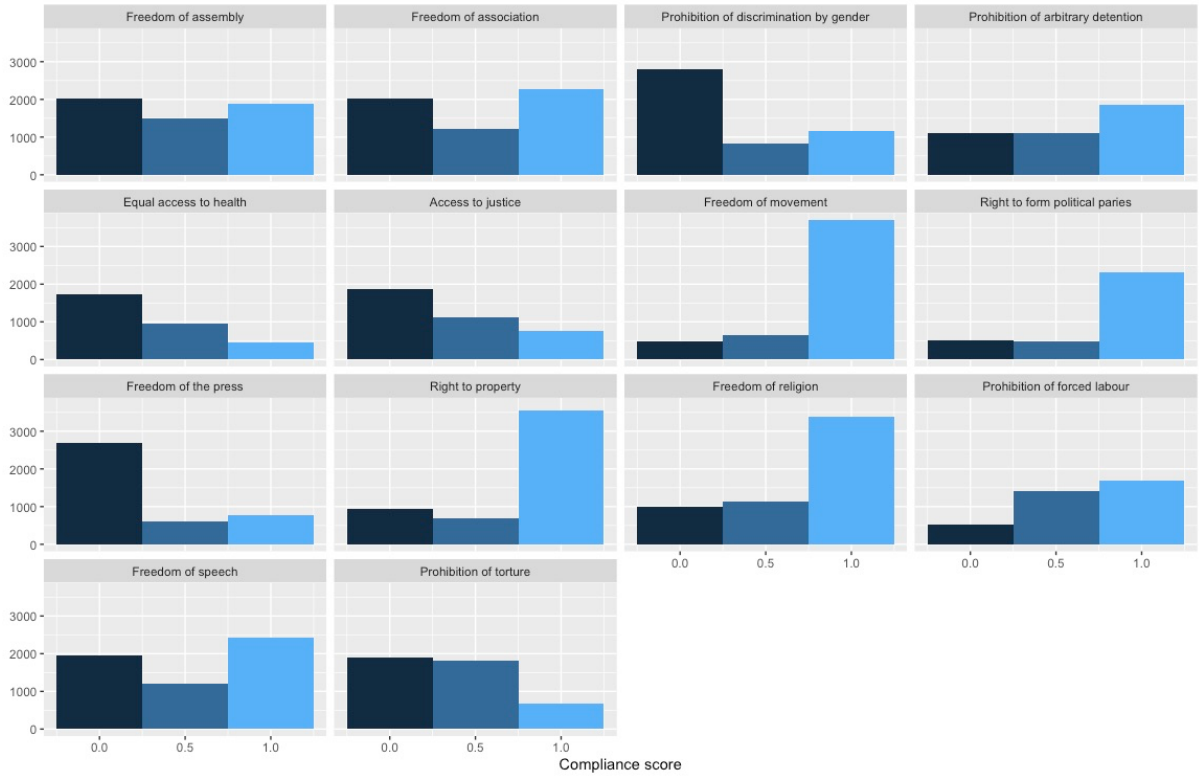


The findings regarding rights-provisions seem to be somewhat in accord with some previous research done into the de jure-de facto gap for civil liberties. Elkins, Ginsburg, and Melton (2009, 29-31; 53-55) provide, in a comparison of de jure and de facto parliamentary power and civil liberties for the year 2006, "[…] some suggestive evidence […]" about how parliamentary provisions better describe the realities of the regime, than do right provisions. They find that the actual level of civil liberties varies in a rather dramatic fashion across countries and need not at all correspond to de jure rights. Their measure of parliamentary powers considers a vast number of provisions when compared to this compliance index, as they used Fish and Kroenig's index spanning 32 dimensions of parliamentary powers. This might account for the slightly different findings, as though the legislature-component indeed performs somewhat better than the rights-component, but they overall perform not to dissimilarly. Their coefficient of variation for all years are 0.5 for the legislature-component and 0.4 for the rights-component. This discrepancy in findings might be brought about by the choice to actively score the gap between de jure and de facto and to assign a non-zero value to country-years with missing provisions, rather than plotting the correlation between de jure and de facto. It might also be a consequence of only having data for a single year.

The component measuring the executive dimension has the highest median score, and the highest scores associated with the entirety of its interquartile range. One point of interest for the dispersion around those four provisions, is Denmark's score of 0. The Danish constitution is an old document, dating back to 1849, and little has changed in the provisions regarding executive power (Møller 2020, 46). This situation seems a good example of one of the theorised effects of constitutional age on compliance; as the constitution grows older interpretations of the text evolve and, at least in common law traditions, this could lead to an increasing divergence between text and legal interpretation thereof (Strauss 1996). In the Danish case, it seems that a combination of stringent amendment rules, political stability and capacity, and enforcement of further rights by for instance EU provisions and the European Convention on Human Rights has disincentivised bringing the text more in line with practice. There is also an understanding, based on nuances in the text, that the monarch is much more restrained and does not actually hold power (Christiansen 2020). These nuances and interpretations are nonetheless lost in the translation of entire documents (without their context) to quantitative datapoints and into compliance scores.

The next figure presents all 14 sub-components from the rights-component. The frequency of values is subtracted by those 0.5's indicating a non-present article. The figure shows how respect for citizens' rights and liberties vary substantially according to different rights. There are also variations in how common the different articles are. For instance, consider the number of observations for the right to form political parties or access to justice against the number of observations for the right to property. Or that of equal access to health against freedom of speech or freedom of religion.

Figure 4.5 The subcomponents from the rights-related component. Observations for missing provision subtracted from the 0.5-group



Looking at the figure, we can see that the two articles, amongst all the rights provisions that are most often violated are those for protection against discrimination, or a guarantee of equality, based on gender, and freedom of the press. Freedom of religion, freedom of movement, right to form political parties and the right to property perform especially well. Wherever there are constitutional articles related to those rights, there is almost only compliance. Prohibition from forced labour is scored as 0.5 almost as often as it is scored as 1. As 0.5 most often translates to a rights provision being "mostly" respected in practice, and in this specific instance is less of an intentional violation by government and more of a failure to

enforce the provision[10], this also seems one of the less violated articles wherever present. The freedoms of assembly, association and speech seem to have the most equal distribution of values indicating compliance or non-compliance.

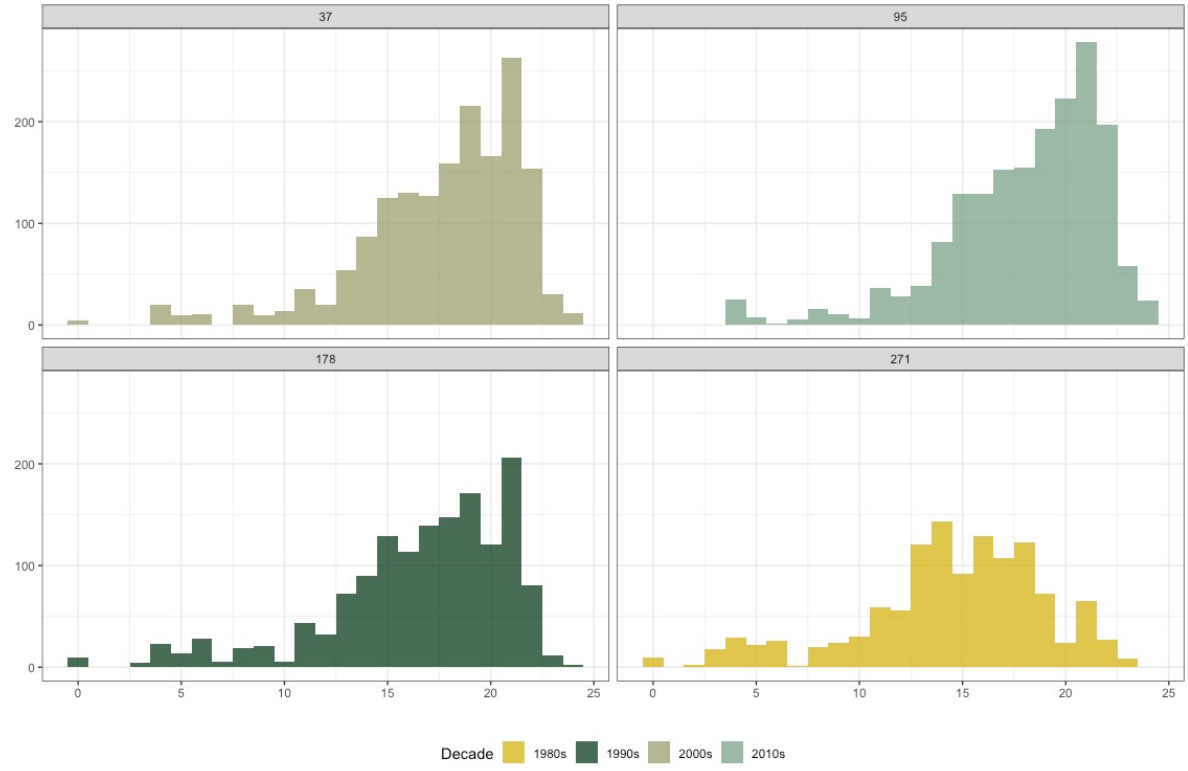
These findings are very much in line with the findings of Law and Versteeg (2013a, 912-913). They find rights associated with gender, freedom from torture and access to justice to be the least respected de facto, whereas freedom of religion, movement and freedom from arbitrary detention are often respected in practice. They also find a pretty even distribution between upheld and violated proportions for freedoms of association, assembly. One major divergence is that they find similar conditions for the right to health. This might be due to different measurement strategies; they have operationalised a right to health as life expectancy at birth, whereas I have measured "[t]o what extent is high quality basic healthcare guaranteed to all […]" (Coppedge et al 2021). As they have combined freedom of the press and expression, and find the compliance rate to be 26.7 %, it looks like the results here do not differ too much from those findings.

4.1.2 With frequencies for provisions present

A first look at the number of provisions present in constitutions over time, seem to indicate an increase for every decade. This matches the expectations due to the witnessing of an increased number of rights provisions included in third wave constitution-making (Elkins, Ginsburg, and Simmons 2013). The relationship does not seem to solely be the cause of increased number of rights provisions recorded as present each year, though, but rather an overall increase through all four components of the index.

[10] See Coppedge et al (2021) for the possible categories related to the de facto presence of forced labour. The third category was for cases where forced labour is considered "[…] infrequent and only found in the criminal underground. It is actively and sincerely opposed by the public authorities."

Figure 4.6 Provisions measured as present, by decade. Group label indication number of missing country-years each decade



As there are some quite large variations within missing country-years distributed by decade, there are, however, some uncertainties in connection to interpreting how the number of provisions constitutions tend to have, have evolved over time. With that uncertainty in mind, there does seem to be a trend towards comprehensiveness over time.

Table 4.2 presents the frequencies for country-years falling into the quartiles of number of provisions present over the four decades measured. The first quartile is for those country-years where the constitution contains fewer than 15 of the examined provisions. The second quartile holds those with 15 to 17 provisions present. The third and fourth quartile encompass 18 to 20 provisions and more than 20 provisions, respectively. In the 1980, 45.57 % had fewer than 15 provisions, whereas this percentage shrunk every following decade. The percentage of country-years where more than 20 provisions were present also increased through time, from 8.35 % to 31 %.

Table 4.2 Number of provisions present, grouped by decade

1980s	1	2	3	4	NA	Total
Frequency	551	338	219	101	271	1480
Percent	45.57	27.96	18.11	8.35		100
1990s	1	2	3	4	NA	Total
Frequency	357	391	439	301	179	1667
Percent	23.99	26.28	29.50	20.23		100
2000s	1	2	3	4	NA	Total
Frequency	286	382	541	458	37	1704
Percent	17.16	22.92	32.45	27.47		100
2010s	1	2	3	4	NA	Total
Frequency	257	411	572	557	95	1892
Percent	14.30	22.87	31.83	31.00		100

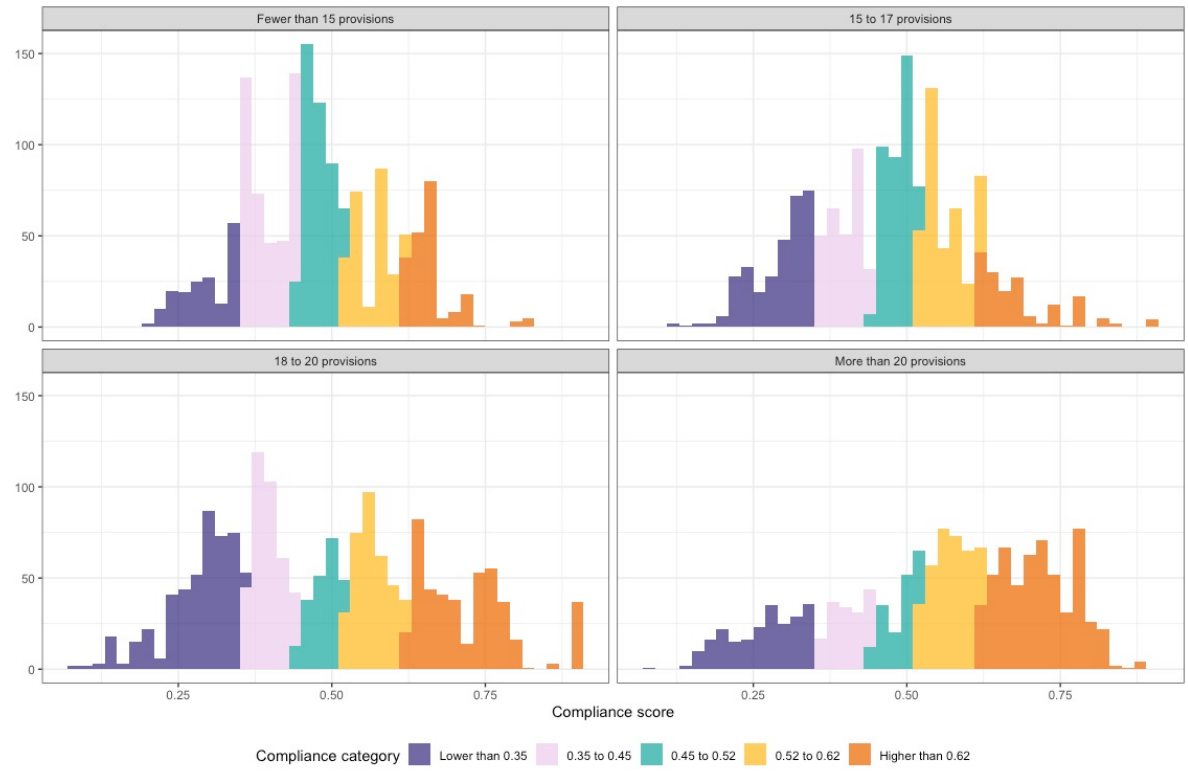
4.2 A look at compliance scores

In this section attention is now turned towards bivariate correlations, such as those between compliance and time, compliance and regional differences, compliance and number of provisions present and finally, compliance and regime type. The visualisations of the data seem to indicate the presence of positive correlation for time, number of articles and for democratic regime. There are also clear regional, overall differences, though with quite a large spread within some of the regions. The regional differences will also be explored by the index-components, as will the regime-differences.

4.2.1 Depending on the number of provisions

Looking at Figure 4.7, there seems to be a rather clear correlation and positive correlation between the number of provisions present in the constitution and compliance-score. In the two groups containing 17 or fewer provisions, more than 70 % of observations fall in the three quintiles falling below the theoretically empty constitution and, therefore, compliance.

Figure 4.7 Compliance, grouped by number of provisions present. Colours indicating quintiles of compliance



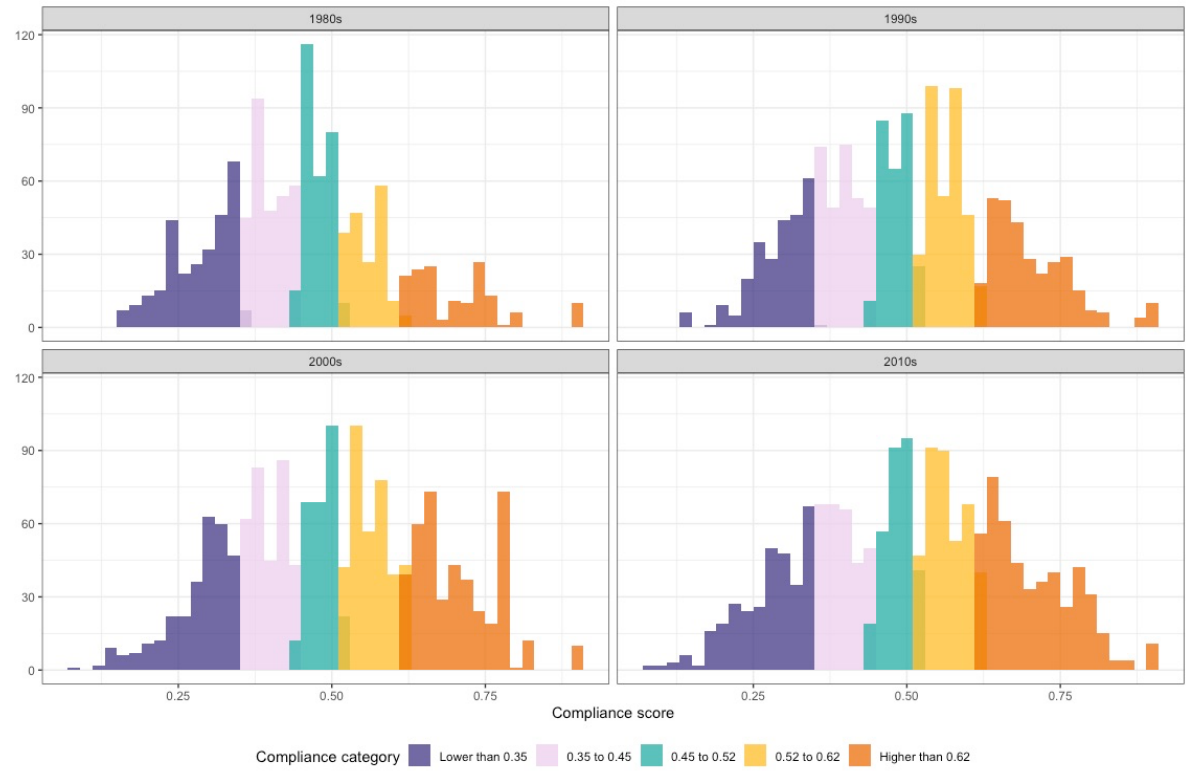
In the group containing 18 to 20 provisions, 58.5 % score below compliance, whereas less than half (40.16 %) of the group containing more than 20 provisions fall below the minimum level of compliance. In this last group, the largest number of country years (36.84 %) score higher than 0.62 on the compliance scale. Those country years scoring lower than 0.35 still amount to the third largest category. The only group containing a lower percentage of country years scoring less than 0.35 is the group containing fewer than 15 provisions (13.09 % versus 17.29 % in the group with the most provisions).

This apparent relationship could be, cautiously as there has been no multivariate analysis, taken as a strengthening of the theory of renegotiation put forth by Elkins, Ginsburg, and Melton (2009); increased specificity and clarity of a constitution might facilitate and incentivise enforcement of the document. While these numbers tell us nothing about clarity, nor about the amount of detail in each provision, which is part of specificity, it does indicate that more rovisions occur alongside higher averages of the compliance index. Specificity is related to longer periods of drafting and higher costs for involved parties. If those involved with the drafting has committed resources into the document, they might be more willing to provide enforcement of it (Elkins, Ginsburg, and Melton 2009, 84-88, 103-106).

4.2.2 Through time

Figure 4.8 demonstrates an apparent relationship between time and respect for constitutional provisions. While 72.04 % of country-years belonged to a category indicating non-compliance in 1980, this shrunk to 55.78 % of country-years in the 1990. In the 2000s and 2010s, respectively, membership in these three groups fell to 53.30 % and 51.53 %. There is, however, a large portion of data missing in the first decade, which is not observable in the below figure. 18.31 % of the 1480 possible observations are missing in the 1980s. This number also drastically shrinks with the passage of time, through 10.68 % and 2.17 % to 5.02 % in the 2010s. The effect of time might then turn out to be less significant than the one indicated by the data and Figure 4.8. The reasons for the shifting membership structure can be theorised to be because of the introductions of many new constitutional systems over the four decades[11], or new measurements uncovering compliance from a previous decade's missing observations. The percentage of observations, missing or not, that fall into the fifth quintile is steadily rising through all decades, from 10.20 % in the 1980s to 25.48 % in the 2010s. What these overall trends might be due to, is another interesting question.

Figure 4.8 Compliance, grouped by decades. Colours indicating quintiles of compliance



[11] The number of observations by the respective decades are 1480, 1666, 1705 and 1892.

4.2.3 Regional differences

Based on previous studies (for instance Law and Versteeg 2013a, 907) there is reason to assume compliance might not only vary by time, but also spatially. To illustrate these relationships are Figures 4.9 through 4.13, where the different boxes are associated with different regions. The grouping by regions is done by politico-geographic association, meaning that some countries find themselves grouped at odds with purely geographic belonging, where the politico-aspect comes from characteristics contributing to regional understanding identified in studies of democratisation (Coppedge et al 2021; Teorell et al 2018). Examples of this is Australia, Cyprus and New Zealand's membership in the Western Europe and North America group, Mongolia's inclusion in the Eastern Europe and Central Asia group, and Turkey inclusion in the Middle East and North Africa region.

As illustrated in Figure 4.9, there are indeed significant differences between the regions. Western Europe and North America is the most compliant region, while the Middle East and North Africa is the least compliant. Asia and the Pacific is the region with the lowest overall scores, after the Middle East and North Africa, with the most dispersion over values. South Korea is the second highest scorer in 2020 at 0.87, only behind Sweden at 0.90, whereas Myanmar is the lowest scorer with 0.18 overall. Sub-Saharan Africa scores substantively better as a whole, the northern part of the continent. The scores vary quite significantly there as well, with Cape Verde at a score of 0.72 (the fifteenth highest score that year) to Sudan at 0.21 (only above Myanmar). There is also considerable range within Latin America and the Caribbean and Eastern Europe and Central Asia, though the regions perform comparatively alike, overall.

The enforcement of provisions related to the executive is shown in Figure 4.10. The decidedly worst-performing category is Western Europe and North America, with a median well below the cut-off for compliance. The best performing region might arguably be Latin America and the Caribbean as they have the entire interquartile range above the cut-off. They do have more outliers than do Eastern Europe and Central Asia, however, even if that region's interquartile range stretches down to 0.5. Sub-Saharan Africa outperforms both the Middle East and North Africa, and Asia and the Pacific. These trends of a high degree of compliance are the inverse of what we see in Figure 4.11. The adherence to constitutional articles related to the independence of the judiciary, meritocracy in public administration as well as the existence of institutions such as the ombudsman capable of executive oversight and electoral commissions,

is overall poor. Only Western Europe and North America has its interquartile range above 0.5. There are outliers from all regions performing within degrees of compliance, safe for the Middle East and North Africa, but the vast majority of countries cannot be considered to follow these provisions in practice.

Figure 4.9 Compliance, grouped by regions

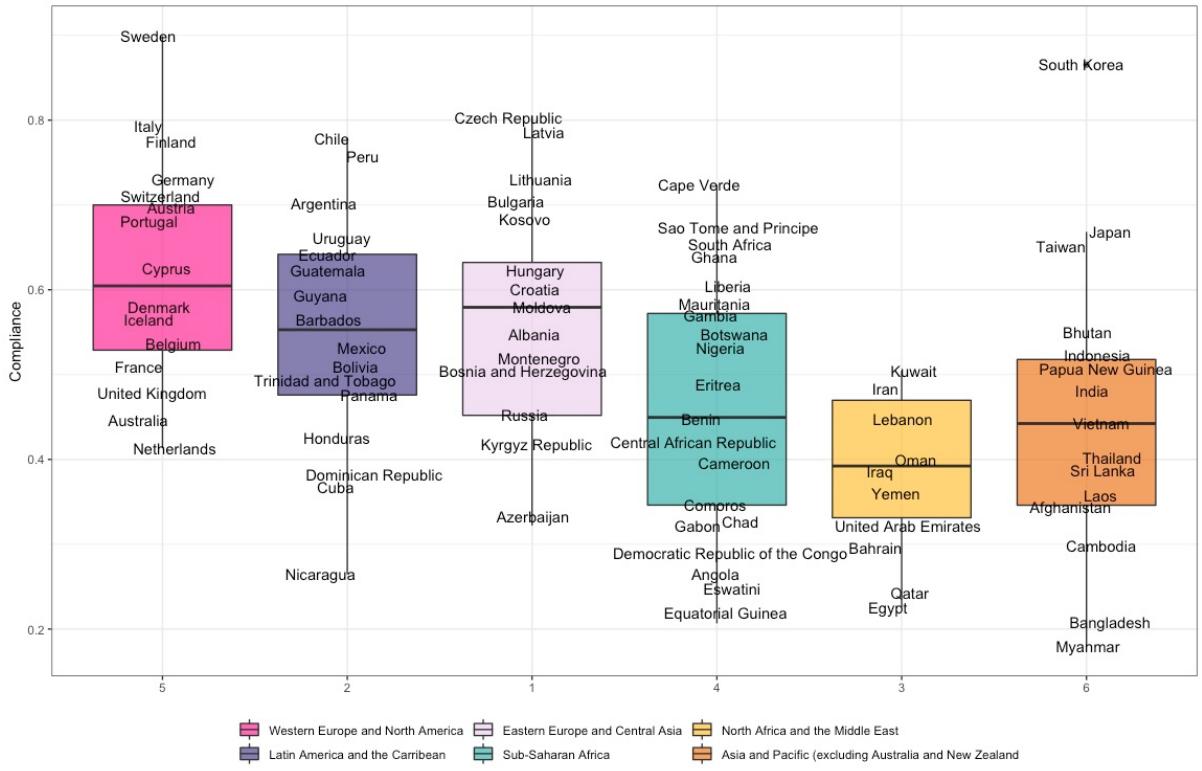


Figure 4.10 Compliance-component for the executive year 2020, grouped by regions

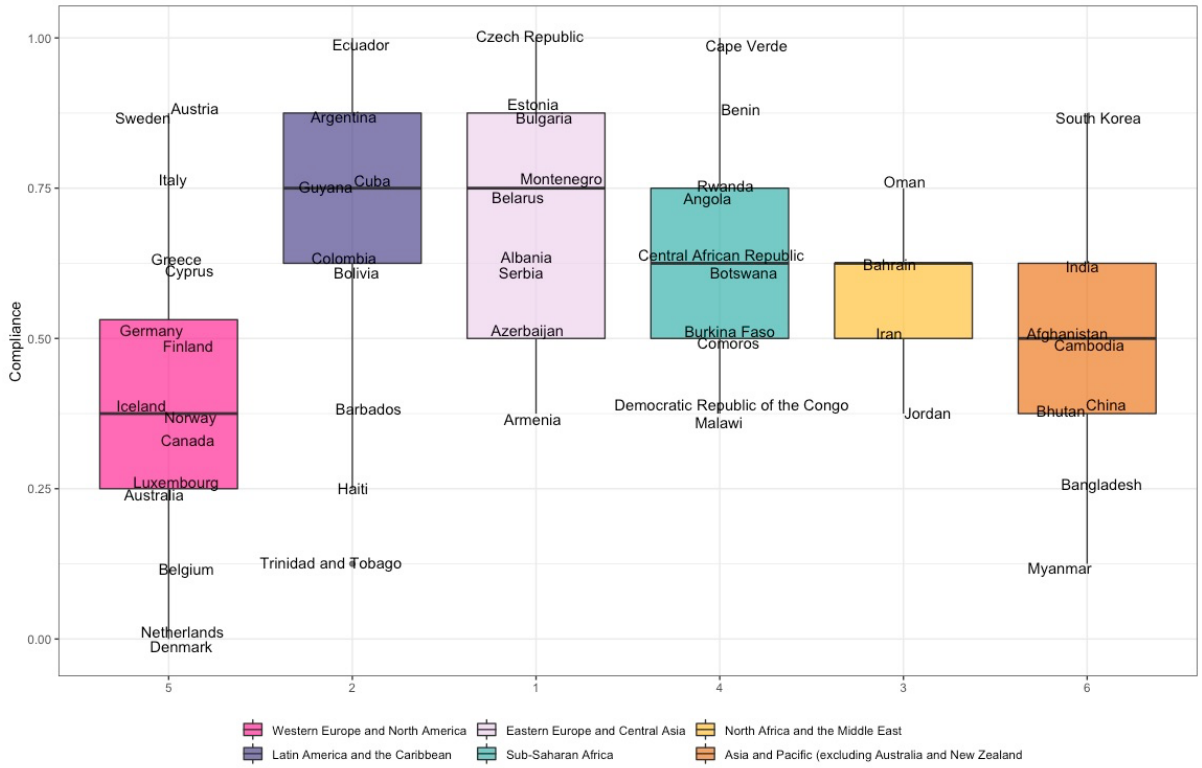


Figure 4.11 Compliance-component for the judiciary and public administration year 2020, grouped by regions

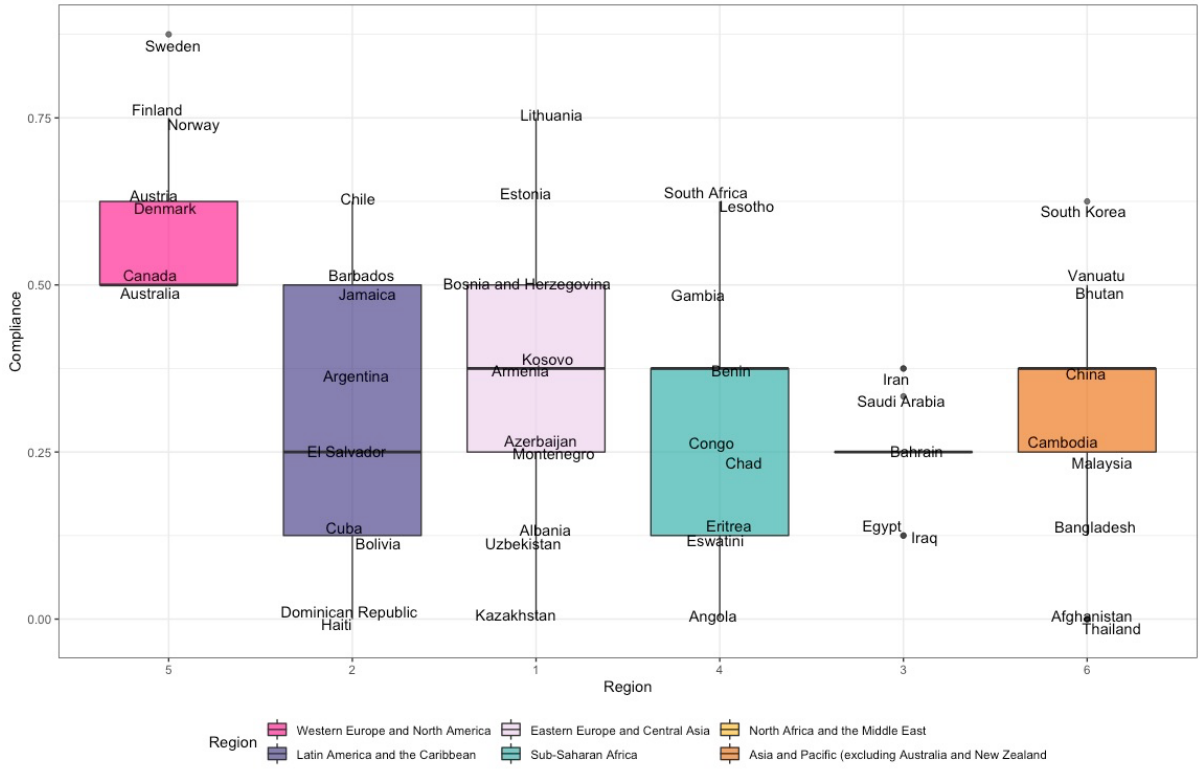
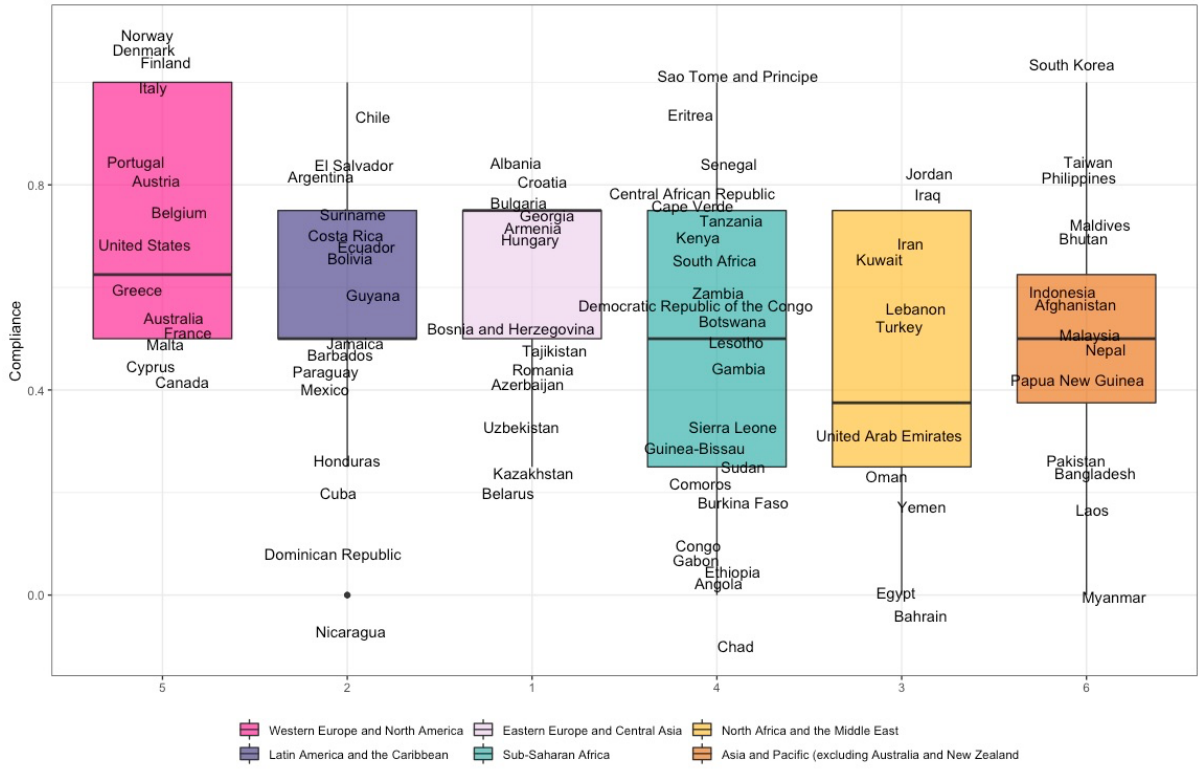


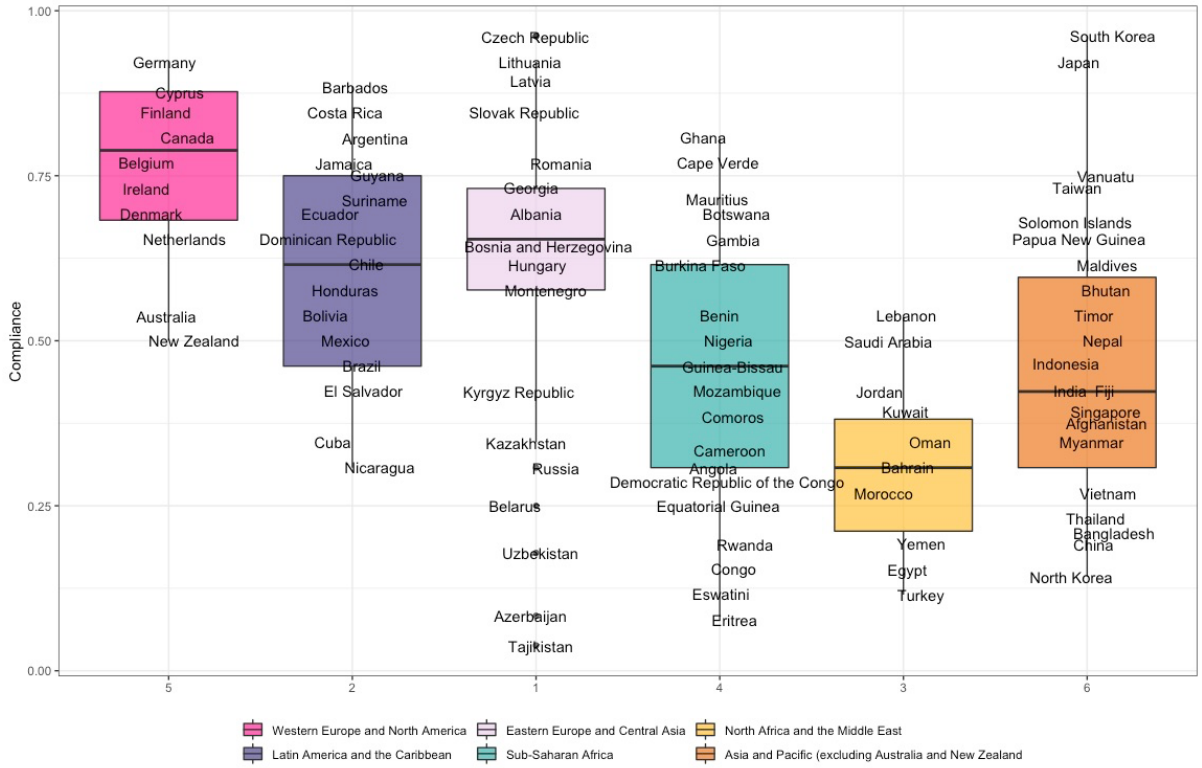
Figure 4.12 maps adherence to the two provisions regarding the legislature onto regional categories. This component sees some mixed results, though they tend to lean towards higher scores rather than lower, overall. There seems to be less of a regional effect for this component. All regions seem to contain countries ranging far away from one another in terms of scores. Only Western Europe and North America do not span far into the non-compliance scores; the poorest performing countries in this region scored 0.5 in 2020. Amongst them are the Netherlands, Iceland, Canada, France, Ireland, Spain and the United Kingdom.

Figure 4.12 Compliance-component for the legislature year 2020, grouped by regions



In terms of keeping with citizens' rights and liberties, as seen in Figure 4.13, there are some marked differences between regions. The least compliant group is North Africa and the Middle East, where only Lebanon passed the threshold of the cut-off point in 2020. The most compliant bunch can be found in the Western Europe and North America group, where only New Zealand with a score of 0.5 and no rights provisions present fell below the cut-off. Latin America and the Caribbean and Eastern Europe and Central Asia following behind. There are, however, rather significant outliers across most regions. Eastern Europe and Central Asia has the most diverse dispersion of countries, with Tajikistan and Czech Republic scoring respectively 0.03 and 0.96 each having 13 provisions. Three of the other top 10 most compliant countries are also found within the Easter Europe and Central Asia group: Latvia, Lithuania, and Estonia.

Figure 4.13 Compliance-component for rights provisions year 2020, grouped by regions



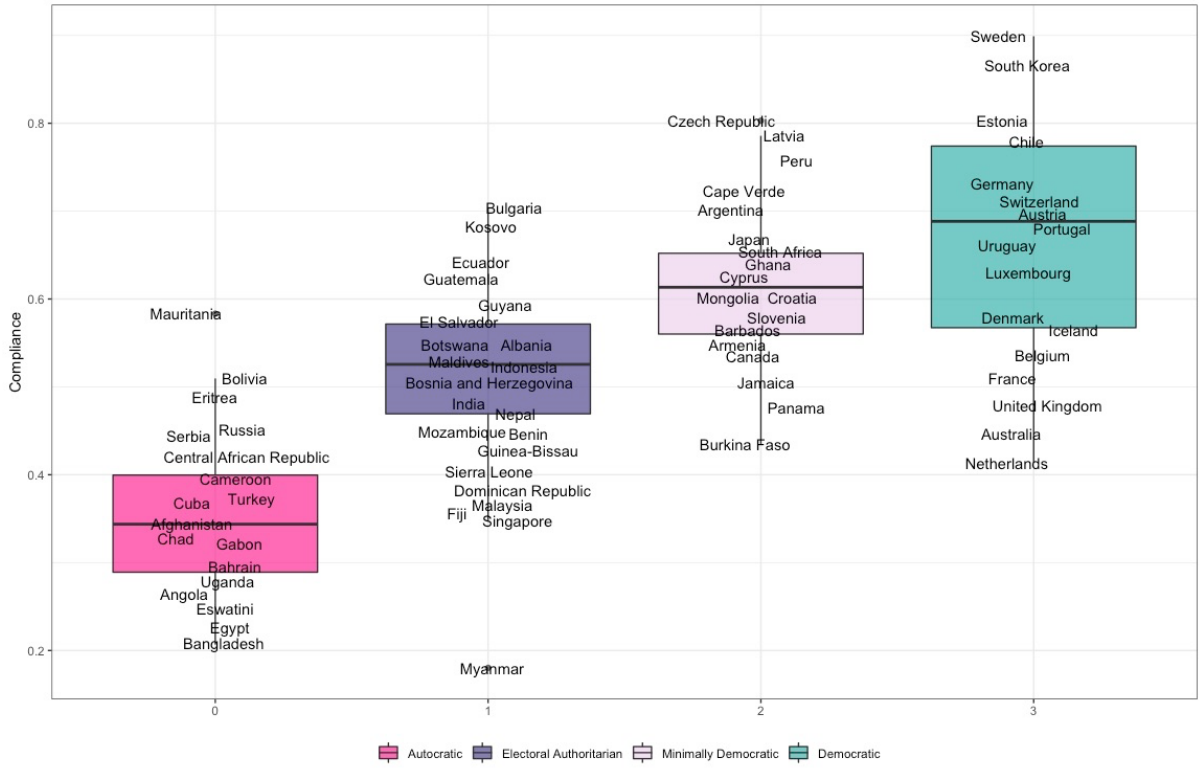
The Asia and Pacific group contain another two from the ten most compliant, namely Japan and South Korea, with respectively 12 and 11 provisions present. North Korea, the groups least compliant state, scores 0.14 for its 10 provisions. Sub-Saharan Africa also sees quite a spread, ranging from Ghana at 0.81 to Eritrea at 0.08. The median scorer in these two groups both fall outside of the compliance-range. The findings of considerable dispersion of scores for rights provisions across Sub-Saharan Africa and Asia Pacific also falls in line with previous findings, so is the dispersion of scores in Latin American and the Caribbean given trends towards more respect for rights provisions since the 1990s (Law and Versteeg 2013a, 911).

4.2.4 Regime type

Democracy is often considered as a determinant of constitutional compliance (X). As we saw in Chapter 2, it is also by some considered as a component of constitutionalism. For those that consider democracy as a vital part of the constitutionalist function, the exploration of the bivariate relationship between these two indicators will perhaps give an indication of cases fulfilling necessary, if not perhaps quite sufficient, conditions for classifying as constitutionalism. Most liberal democratic countries score well above the cut-off point for constitutionalism (roughly 0.52 for these data), though there are those that score below it as

well (France, the United Kingdom, Australia, and the Netherlands). The same holds true for minimally democratic countries, where most score above the cut-off, but Panama, Jamaica and Burkina Faso fall below.

Figure 4.14 Compliance by regime type, year 2020



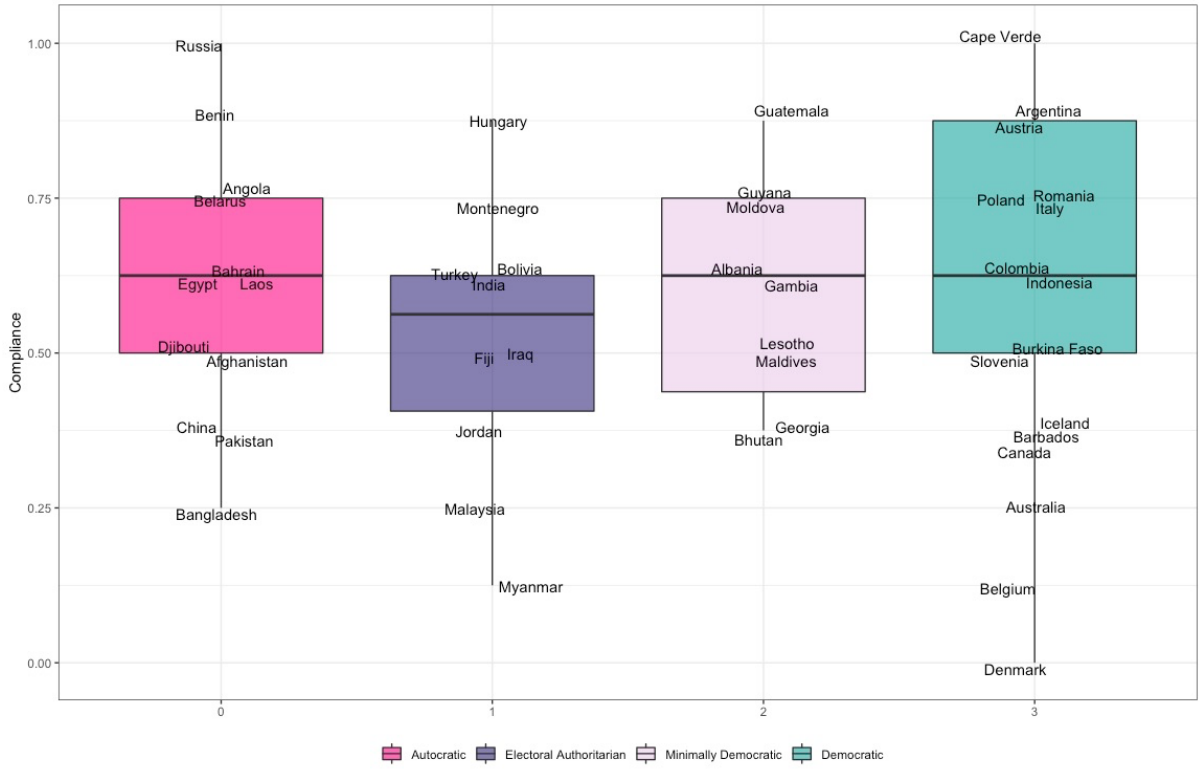
If one is inclined towards conceptualisations of constitutionalism that exclude democracy, the relationship portrayed is still of interest. Liberal democracy, as understood here[12], could be correlated with compliance due to the conceptualisation of liberal democracies as countries that enforce constitutional protections of civil liberties, have a strong rule of law, independent judiciary, as well as checks and balances seeking to minimise governmental power (Coppedge et al 2021). This sounds, in many ways, like an operationalisation of a de facto measurement for restraining government power. It is not far from the conceptualisation of the de facto effects a constitution must have to fulfil the compliance-condition for constitutionalism. Fitting the values onto a framework of electoral democracy instead (Coppegde et al 2021)[13], we see the same general trends, but with some adjustments to the boxes position. The figure can be found in the Appendix. For the remainder of the figures in this section, liberal democracy has been

[12] The measure used is the V-Dem Liberal Democracy Index.

[13] The measure used is the V-Dem Polyarchy Index.

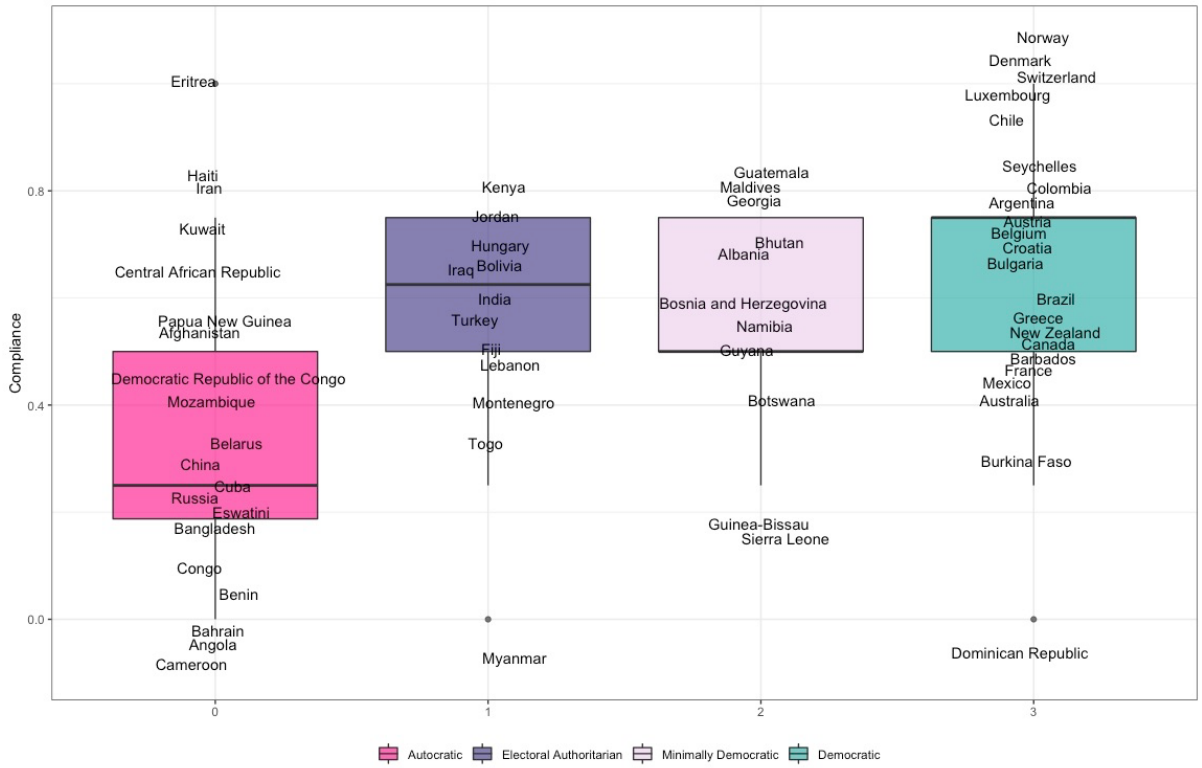
exchanged for electoral democracy to limit the potential for conceptual overlap of the indicators.

Figure 4.15 Compliance-component for the executive year 2020, grouped by regime type



The executive-component shows a fascinating correlation, where autocratic regimes by no standard can be considered the category with the lowest level of compliance. This might serve as further empirical indications that the theories regarding constitutional functions and whether some form of constitutionalism might exist even in non-democratic states. There is a general need, across regime types, for coordination among actors, which could arguably be what we are witnessing in Figure 4.15. Ginsburg and Simpser (2013, 2) argue the function of coordination provide means through which authoritarian constitutions ought not always be considered "sham" documents. In order to provide coordination among elites, there needs to be some form of compliance, or a function of existing as operating manuals as described by Przeworksi (2013). "While authoritarians and democrats may differ in the precise character of the commitments they wish to undertake, the basic modality of entrenching certain policies to enhance credibility may be useful to all leaders" (Ginsburg and Simpser 2013, 4).

Figure 4.16 Compliance-component for the legislature year 2020, grouped by regime type



When it comes to the legislature-component, however, the autocratic group of countries have the lowest level of adherence to the provisions. There are a number of outliers, resulting in a few rather high scores for countries such as Eritrea, Iran, Thailand, Iraq, Mauritania and Bolivia, to mention a few. The most interesting pattern that occurs in this component is arguably the identical positioning of the boxes for the other three levels of democratisation.

Figure 4.17 shows a correlation more along the lines of the overall compliance scores across regime types. As do Figure 4.18. There are still quite a few outliers, and some very heavy ones in the democratic regime category, with Dominican Republic at 0 and Sweden with the overall highest score at 0.875. The rights-component is structured in a similar way, but with overall much higher scores. Whereas both categories of democracy had their interquartile range below 0.5 in the judiciary-and-administration-component, in Figure 4.18 they are both much higher than both 0.5 and the cut-off mark. A few outliers fall below compliance with citizens' rights in the democratic regimes in 2020, such as Brazil, El Salvador, Guinea-Bisseau, Sri Lanka, Nepal, Nigeria, and Indonesia. There are also those from the two autocratic categories that make it over the cut-off score, even if most countries belonging to these two regime types are non-compliant in terms of rights provisions. Among these are Gabon, Lebanon, Bolivia, and Tanzania, as well as some outliers visible in the figure.

Figure 4.17 Compliance-component for the judiciary and public administration year 2020, grouped by regime type

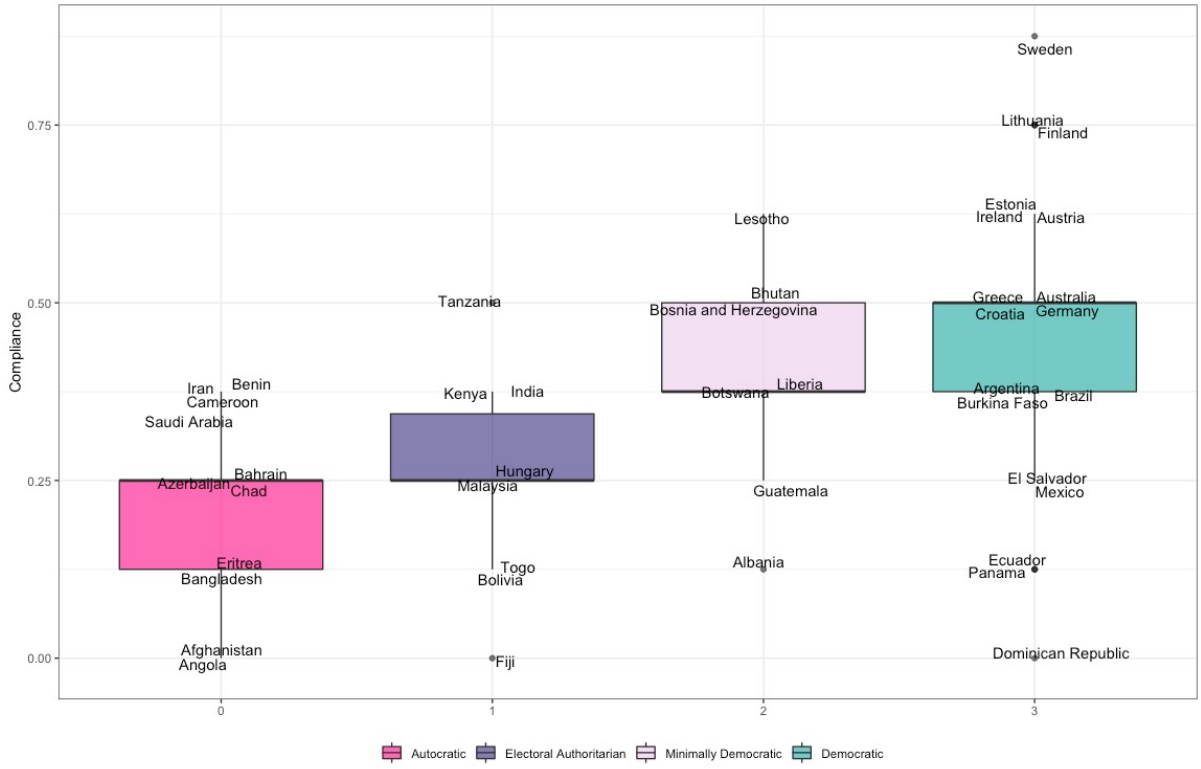
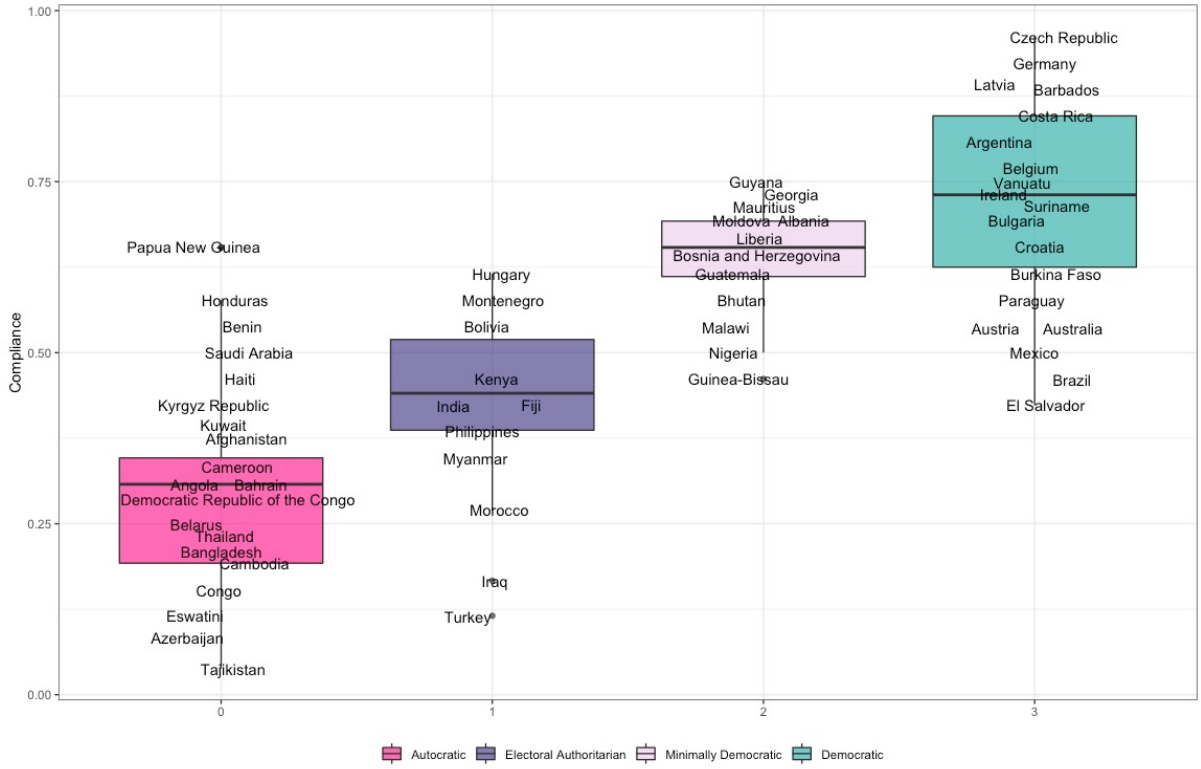


Figure 4.18 Compliance-component for rights provisions year 2020, grouped by regime type



Democracy can be seen to correlate with compliance overall. It is often considered as a predictor of compliance, though this relationship might need some revision if the tendencies seen in Figure 4.16 and especially Figure 4.15 can be believed to be more than merely correlational. The constitution might in turn be a facilitator of constitutional democracies; securing rights which enable an independent media and civil society, and then mitigating coordination problems amongst those seeking to enforce it (Elkins, Ginsburg, and Melton 2009, 77). This is not to say, however, that all non-democratic regimes necessarily violate their constitutions. Mauritania, for instance, scores 0.58 while being considered as a closed autocracy. The median score for countries belonging to the electoral autocracy-category will also score above the cut-off, though barely.

There can be considerate costs associated with making empty promises, and other functions constitutions perform, such as coordination, depend on some level of compliance with the higher law (Law and Versteeg 2013b, 166). In cases where authoritarian elites write constitutions containing many rights, they are at once depriving themselves of a coordination device, while potentially giving one to regime opponents. Even among authoritarian countries there are different approaches taken to the function a constitution is thought to perform, and evidence suggests military and monarchical authoritarian regimes tend to provide fewer rights provisions than do civilian and party-based authoritarian countries (Law and Versteeg 2013, 186b). The general spread *within* the categories indicate that these relationships and internal variations might be of interest for further research with an expanded and improved upon compliance-index.

5 Analysis

In the previous chapter some preliminary analysis of the compliance-index and some potential bivariate relations were visually presented and commented upon. This chapter seeks to further explore some of these relationships, as well as to identify other predictors of compliance. The following section will explore determinants of constitutional compliance, among them democratic level, economic development, comprehensiveness of the constitution, country-level variations, constitutional age, level of civil society participation, population size and judicial independence. The second section concerns the method of analysis, which is a multilevel model, and the underlying assumptions. Lastly follows a discussion of the results from the model specification.

The choice of method needs be anchored in the research question, which for the purposes of this chapter will be *what are likely predictors of constitutional compliance?* To account for variances between countries and for the longitudinal structure of the data, a multilevel method is applied. The choice of method is further explained in section 5.2.

5.1 Determinants of constitutional compliance

The chapter's analysis concerns the potential relationship between the compliance measure generated and an increased number of provisions present in the constitutional text. There will obviously be uncertainties associated with any regression results, as the compliance index cannot be seen as a complete measure of constitutional compliance. There are many indicators excluded from the creation of the compliance measure, especially for provisions that do not involve the rights of the citizens. The index does, however, give us a quantitative view into some aspects or parts of constitutional compliance, and any findings might provide a direction for further inquiry with ameliorated measurements.

Many of the determinants listed in this section come from the literature on human rights violations by states, or non-compliance as concerns rights provisions. This is in large part due to the direction previous quantitative research has taken in this field. As many of the subcomponents of the index assess the performance of rights provisions, it does not seem unreasonable to expect previously identified determinants to play some role in predicting

compliance scores either. Some of the relationships might not, however, apply the same way to the structural provisions. It is uncertain if the so-called "usual suspects" (see for instance Chilton and Versteeg 2014, 582; Keith, Tate, and Poe 2009, 648) perform as they would be expected to, had the compliance index only included rights provisions. The analysis might give us an interesting glimpse into potential associations between compliance for increased parts of the constitution and determinants of de facto rights protection.

5.1.1 Comprehensiveness of the constitution

The comprehensiveness of the constitution is here measured by the number of provisions present in any given country for any given year. The variable is a result from the development of the compliance index, and its' data therefore from Elkins and Ginsburg (2021). Elkins, Ginsburg, and Melton (2009) develop a theory on the role of specificity for constitutional survival, comprised of both the number of topics (scope) and amount of detail included. Their theory posited that larger degree of specificity typically meant a longer constitution-making process, which reflected to some degree different actors' commitment to the constitutional bargain. A constitution of more scope would mitigate problems of unsolved coordination or hidden information, and be overall more self-enforcing than less specific documents (Elkins, Ginsburg, and Melton 2009, 84-87, 103-106).

Law and Versteeg (2013a, 929) and Metelska-Szaniawska (2021, 189) find, on the other hand, that comprehensiveness of the constitution is associated with fewer rights upheld in practice. The operationalisation of comprehensiveness utilised in those studies are a measure of the number of rights a constitution includes, but also a distinction between which countries promises "generic rights" and which countries also promise "esoteric rights" (Law and Versteeg 2013a, 927). For this measure of constitutional comprehensiveness, level of democracy, GDP per capita and constitutional age were considered as determinants (Metelska-Szaniawska 2021, 187).

The prediction for the relationship might then both be negative and positive. Both seem very plausible. More provisions promised gives more occasion for failure, but few provisions promised can make high scores less attainable. Given that my measure is an imperfect capture of both these versions of comprehensiveness, it might not match either. The results from the

bivariate, visual analysis from the previous chapter, constitutional provisions seemed positively correlated with a higher degree of compliance. Yet, to achieve scores approximating 0, a country would need to promise more than not.

5.1.2 Time

There are two variables used to capture the measurement of time in the analysis. One is the indicator for year in which an observation takes place. Elkins, Ginsburg, and Melton (2016, 258) find that protection of rights has decreased over their time-period of study, which is the 1980s until the beginning of the 2010s. They theorise this is due to increased measurement, rather than increased state repression.

The other variable measuring time is the age of the constitution. The indicator is created from the CCP data (Elkins and Ginsburg 2021), where the year of promulgation for a constitutional system is subtracted from the year of any given observation. There are inevitably problems arising from attempting to measure constitutional age, with regards to what qualifies as resetting the number to 0. Amendments can sometimes radically change a document, in such a way that it ought to be classified as a new system. Likewise, replacements can provide startlingly similar documents to their predecessors. There seems to be at least evidence suggesting amendments cover more of the same topics (97 % match) than do replacements (81 % match), even if there is less knowledge on substantial differences within those topics (Elkins, Ginsburg, and Melton 2009, 55-59). As such, the operationalisation of constitutional age is time since replacement in this analysis. There are some country-years where there is no year of promulgation for the constitutional system. For some countries this seems to be because there is no constitution in force those years. For others it seems this is due to the suspension of the constitution.

Two countries had values indicating it had a non-suspended constitution in place: Norway and the United Kingdom. The dataset does not go back further than 1789, so there is no possibility of a value indicating a new constitutional system for the British as the Magna Carta dates to the 1200s. The introduction of age in terms of the British constitutional system also seems daunting, both for the purposes of statistical modelling and for classifying such an exceptional case in the world of constitutions. While the Norwegian constitution presents an outlier as one

of the oldest written constitutions still in force, the British system presents an absolute extreme outlier. The Norwegian constitution, unlike the British, is fully contained within the dataset. The first year of the constitutional system was coded as an amendment, rather than a new constitution. This is probably due to coding rules, as it is specified that the latest event in any given year is the one coded, and there was indeed constitutional amendment in the first year of the Norwegian system (Stortinget 2020). There is, however, another variable, indicating whether multiple events have taken place any given year. This is the case for Norway in 1814. This country-year is also the only for which amendment and multiple events have been coded simultaneously. I therefore chose to add the information of this constitution to the variable measuring age, while feeling confident that this is a relatively isolated case.

The theorised relationship between age and constitutional compliance is somewhat involved. Constitutional age might contribute to a decrease in compliance, as interpretations of the text evolve (Strauss 2010), as found in some research (Metelska-Szaniawska 2021, 188-189), while no significant relationship was found by others (Law and Versteeg 2013b, 929). Another way to see the relationship is to assume it has the opposite effect. Instead of a decrease, it might increase in fit as it grows into its aspirations. The consequence of increasing age, found by Elkins, Ginsburg, and Melton (2016), is a maturation effect and especially so in autocratic regimes. There seems to be two predicted outcomes for the variable of age then, each in the opposite direction.

5.1.3 Space

One huge indicator of variance in compliance scores is spatial context. Different social, political, historical, and cultural contexts dictate both the contents of the constitution and the de facto circumstances, giving rise to very different scores among countries as seen in some of the graphic representations of the data in the previous chapter. The effects at play in a specific country is, however, not of particular interest. Therefore, these differences will be accounted for by allowing for within-country variance in the model. Data on countries come from the CCP and V-Dem datasets (Elkins and Ginsburg 2021; Coppedge et al 2021).

5.1.4 Democracy

"[…] studies have consistently found a link between democracy and government violence" (Chilton and Versteeg 2016, 582). Poe and Tate (1994) find a strong association between democracy and a lower level of rights repression by governments. They theorise the effects is due to the very real possibility of ouster of any political leader seeking to curtail or abuse human rights. Democracy is consistently included in analyses of the effect of and on rights provisions (see for instance Chilton and Versteeg 2016; Davenport 1996; Fox and Flores 2009; Metelska-Szaniawska and Lewkowicz 2021). The relationship between the compliance index and democracy is expected to be positive, primarily due to the findings in previous research concerned with predictors of de facto rights protection but also due to the bivariate visualisations from Chapter 4. The data for the democracy variable is the Electoral Democracy Index from the V-Dem dataset (Coppedge et al 2021).

5.1.5 Civil society engagement

The presence of a strong and organized civil society is often theorised to perform a function of enforcement. This is due to non-governmental organisations capabilities, alongside accessible communication channels, of providing information to the general public of breaches to the constitutional bargain. It is also in their own self-interest to work towards the ensuring of rights that enable their own functioning, and they might do so both through pre-emptive actions or reactively (Chilton and Versteeg 2016). Their presence thus function as a disincentive for governments preferring non-compliance (Ben-Bassat and Dahan, 170). They perform functions such as those of other accountability structures, such as the judiciary or independent agencies.

In order to measure a strong civil society, the civil society participation index from V-Dem is used. The index "[…] is designed to provide a measure of a robust civil society, understood as one that enjoys autonomy from the state and in which citizens freely and actively pursue their political and civic goals, however conceived" (Coppedge et al 2021). The effect of civil society engagement on compliance is, through the mechanisms mentioned above, expected to be positive.

5.1.6 Economic development and population size

From studies on human rights (see for instance Poe and Tate 1994), there have been established links between wealthier countries and fewer human rights abuses, though the effects were found to be more modest than those of democracy and conflict. The lack of economic growth is thought to lead to unstable regimes, more prone to rights violations. Growth, however, might alleviate scarcity of resources, leading to a lesser need for repression. It might also, if it is of a rapid nature, create instability which again leads to more repression (Poe and Tate 1994, 857-858). It could also be argued that the economy and constitutional compliance has a reciprocal effect, wherein compliance further incentivises investments and growth, regardless of the level of democracy in the regime (Albertus and Menaldo 2014).

The expected effects are, because a measure of development rather than growth is used, a positive effect on compliance. Economic development is here measured by an indicator for GDP per capita. The indicator is accessed through the V-Dem dataset, and uses data from the Maddison Project Database (Bolt and van Zanden 2020; Bolt et al 2014; Coppedge et al 2021).

Population size is expected to correlate negatively with compliance. It is theorised that larger populations might create more opportunities for or cases of repression so to speak, and a larger population is thought to increase the burden on national resources which potentially destabilise political rule and therefore incentivises violence by the state (Poe and Tate 1994, 857). Data on population size is from World Bank Development Indicators (2019), accessed through the V-Dem dataset (Coppedge et al 2021).

5.1.7 Conflicts

Involvement in violent conflicts is now seen as one of the "usual suspects" when it comes to affecting regime repression, and it is found consistently as a significant predicator of violence by the state (Hill and Jones 2014). Violent conflicts may also by many of the same mechanisms, at least those relating to capacity, be thought to influence compliance with other aspects of the constitution as these forms of violence often pose threats to and exerts pressure upon those occupying government offices. The indicator of conflict is a measure of civil war from Haber and Menaldo (2011), accessed through the V-Dem dataset (Coppedge et al 2021). Civil war is expected to have a negative effect on constitutional compliance.

5.1.8 Judicial independence and government adherence to judicial decisions

Judicial independence is widely considered to be a determinant of a small gap between de jure and de facto (Chilton and Versteeg 2016, 582). Metelska-Szaniawska and Lewkowicz (2021) find that relatively higher judicial independence affects, in a positive manner, the de facto enforcement of rights. An independent judiciary provides the capabilities of enforcing constitutional articles, and therefore operates as a condition for compliance. It requires implementation of judicial decisions, regardless of whether these decisions are in the interests of the implementors (Voigt, Gutmann, and Feld 2015). It should be noted that the judiciary is not the only actor capable of solving the commitment problem on behalf of the government; independent agencies, such as public prosecutors, can also fulfil such functions (Voigt, Gutmann, and Feld 2015, 199).

Judicial independence is both sought after by political leaders, as the presence of such solves problems of credibility on behalf of the government, but it is also unwanted when the wishes of political leaders run contra to the rulings of the courts (Feld and Voigt 2003). The expected effect of a higher level of judicial independence is therefore a higher level of compliance with the constitution. It might also be of note that by acting as an enforcer of the constitution, de facto judicial independence is also associated with facilitating economic growth (Feld and Voigt 2003; Voigt, Gutmann, and Feld 2015), which is another one of the determinants of rights compliance, suggesting these processes feed into each other.

In order to control for level of judicial independence, two indicators are included in the analysis. These are indicators of expert measurements for high court independence and compliance with the high court (Coppedge et al 2021). The indicator for high court independence asks for how often the court rules in favour of the government in cases deemed important to the government, but only when these rulings are thought to go against the court's own view of the law The indicator for compliance with the high court covers the other side of this relationship by asking how often "[…] the government complies with important decisions of the high court with which it disagrees" (Coppedge et al 2021).

5.2 Multilevel timeseries cross-section analysis

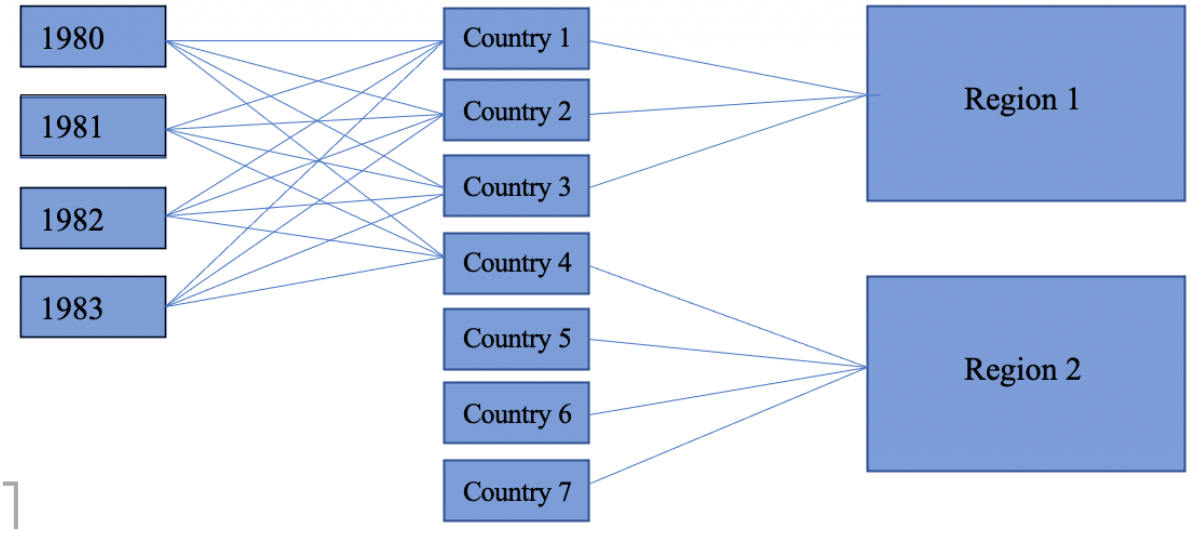
Multilevel analysis, sometimes called mixed models or hierarchical models, refer to a form of analysis aiming to take context of different levels into account (Robson and Pevalin 2016, 6). It allows us to study effects that vary within and across groups (Demidenko 2013, 4; Gelman and Hill 2007, 6; Western 1998, 1233). Classical assumptions in statistics involve observations being drawn from the same general population, being both independent of one another and identically distributed. Multilevel analysis provides methods of analysis where these assumptions cannot be met (Demidenko 2013, 1). They are typically associated with less stringent demands, such as being capable of handling partially missing data, unbalanced data, and nonstandard intervals in time between measurements (Brown 2018, 496).

Analyses assuming multilevel structures can be used for multiple observations nested within a single object, making it a relevant form of analysis for longitudinal data where one is concerned with structure and predictors' change over time (Luke 2004, 62-63). An assumption of longitudinal data nested hierarchically allows for simultaneous modelling of change within an individual and differences in change within time across individuals (Finch, Bolin and Kelley 2014, 100). For longitudinal pooled data, multilevel models can conceive of cross-sections as nested within time-periods, or as time-periods nested within cross-sections (Raffalovich and Chung 2014, 211). Multilevel analysis in terms of time, however, can be considered as non-nested, or crossed, when individual observations are nested within levels that are not of a clear hierarchical nature (West, Welch and Galecki 2014, 369). Examples of these non-nested structures include country-year within both country and year in timeseries cross-sectional data (Gelman and Hill 2007, 243-244).

The structure of the data then seems applicable to both a non-nested, or crossed, multilevel model and to a hierarchical nesting of years within states. The assignment of variables to different levels of a structure can, in reality, be somewhat unclear (Hox 2010, 7). What structure is applied to the data seems more indicative of researchers' perceptions and intentions than of any inherent properties in the data (Raffalovich and Chung 2014, 211). Applying an understanding of non-nested or crossed levels, we would expect the years and states measured to have independent, linear effects on compliance. For a nested level-structure, we might expect that the effect observations within time have on compliance is completely dependent on states. A nested structure sees country-year observations within country, thereby accounting for the passing of time within level two units, whereas the non-nested structure sees country-year as nested within two different level two groups simultaneously.

Figure 5.1 illustrates an argument for why time itself ought not to be considered as nested. Each value of the time-variable occurs within every value of the country-variable. The relationship is crossed, rather than nested. Each value of the country-variable, however, corresponds to only one value on the variable identifying regions. This is a nested structure, and as such ought to be modelled that way if region was included in the analysis.

Figure 5.1 Illustration of nesting structures



There are theoretical, empirical as well as statistical reasons for applying an understanding of the data as structured within multiple groups (Luke 2004; Robson and Pevalin 2016). Multilevel analysis is better equipped for modelling the effects of predictors existing at different levels (Gelman and Hill 2007, 7; Robson and Pevalin 2016, 7; Finch, Bolin and Kelley 2014, 100). If we can reasonably assume a multilevel structure among the dependent variable and relevant predictors and thereafter apply multilevel analysis, we might avoid errors such as ecological or atomical fallacies (Robson and Pevalin 2016, 4-6).

From a theoretical standpoint, the data should be considered as multilevel due to the assumed varying effects within the passage of time or within differing countries. The empirical data indicates a variance within time and within countries for the index, as shown in the previous chapter, which further justifies considering the data as a multilevel structure. The reasons for choosing a form of analysis considering the multilevel structure of the data are, as shown, manifold. One of the most important statistical reasons is the assumption of independence for

OLS regression. "This assumption essentially means that there are no relationships among individuals in the sample for the dependent variable *once the independent variables in the analysis are accounted for*" (Finch, Bolin and Kelley 2014, 23). There is no reason to assume compliance scores are unaffected by grouping within time or country. Scores from one year will likely be correlated to scores from the previous or following year. Variables other than the dependent are also likely to be influenced by country or time groupings. Population size will most likely be strongly correlated within a country, as will GDP per capita. We can therefore assume a violation of this assumption and resulting incorrect estimates of standard errors or spurious significant results (Hox 2010, 3). Effects of particular predictors may also vary in their effect in different contexts, which is something multilevel analysis allows for (Robson and Pevalin 2016, 8-16).

5.2.1 Specifying the model

In order to address the statistical basis for analysing the data using a multilevel structure an empty model has been tested for its intraclass correlation coefficient (ICC), and models with different specifications in terms of random and fixed effects. The first model has both year and country as varying intercepts. The second model has year as a fixed effect, whereas the intercept for country is allowed variation. The other predictors will be held as fixed terms both models.

Empty models with varying intercepts for year and for country, for only year and for only country gave different indications that within-time variance might or might not be a good explanatory variable. Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for the model that included both random intercepts indicated that to be the preferred model, as they were lower than for the other two empty models (Hox 2010, 92-93). The likelihood ratio test of the same model was statistically significant at the 0.001 level, also indicating that there are indeed varying effects of both time and space at play. The ICC for the model is at 0.811. When checking the other two models, however, country-varying intercepts had an ICC of 0.797, while year-varying intercepts had an ICC of 0.015.

This is an argument against the inclusion of time as a level two grouping. The alternative would be to simply nest country-years in countries, which does account for the effect of time within

states. However, as these are mostly considered guidelines rather than absolute rules, because the AIC and BIC results indicate a preference for models with year modelled as a varying effect, and because of the recommendations of Barr et al (2013), time is kept as a level two group. Barr et al (2013) argue that identifying maximal random effects structures should be considered as best practices. Models lacking random effects when they can be theoretically presumed to be there will through unaccounted-for variation reduce the power of the tests on the effect of interest. Exclusion of random slopes can lead to a confounding of effects, leading to increased risk of Type I error (Barr et al 2013, 261). In order to truly follow this advice, the addition of random slopes into the models might have seemed like a good choice, but as there are no clear theoretical reasons to expect varying slopes for these relationships this has not been pursued.

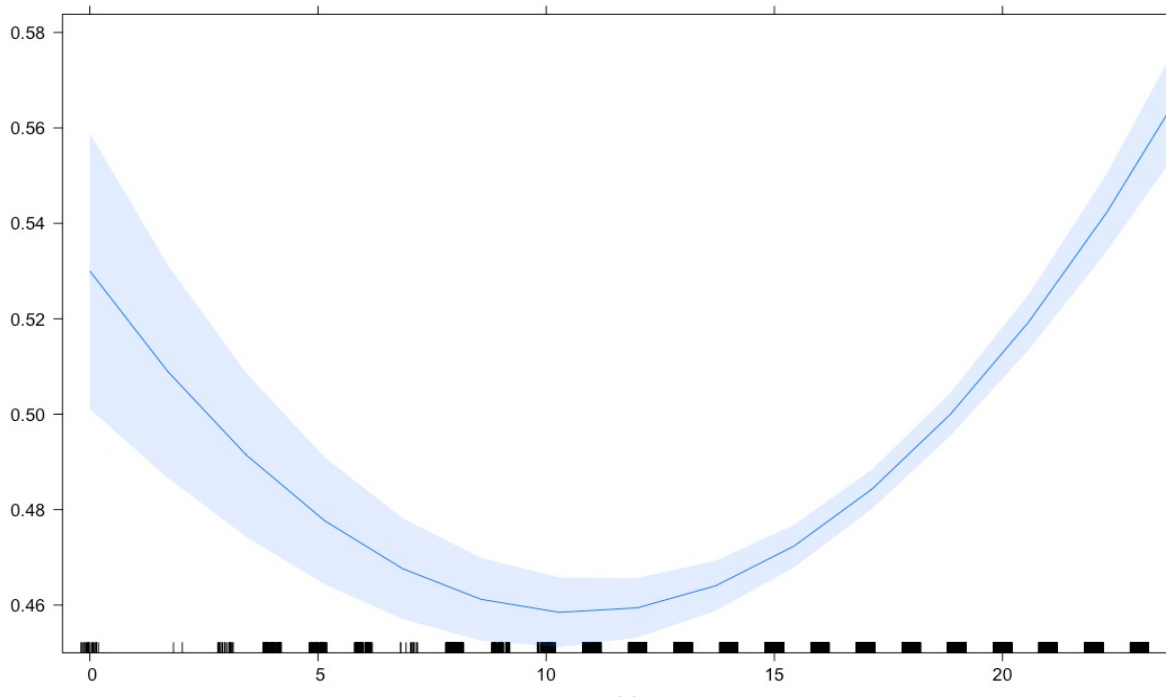
To compare different models and decide on final specifications, AIC, BIC and likelihood ratio test are used, as deviance is preferred for testing nested-models and AIC and BIC is preferred for testing models that are not nested (Hox 2010, 47-50). One model is fitted with the lme4-package and the other with the nlme-package in R, with a further discussion on reasons for this in the section on regression assumptions. The two models that were specified with this method are contrasted to two fully specified models, one with only varying intercept for country and another non-nested one, containing all the theoretically chosen predictors discussed in the first section of the chapter.

While the specification of fixed effects was done with maximum likelihood estimation (ML), the rest of the specifications were done with restricted maximum likelihood estimation (REML). REML include only variance components in the likelihood function, while the regression coefficients are measured in a second step. REML estimate variance components after excluding fixed effects from the calculation, which leads to less bias than for ML (Hox 2010, 41). ML is used for specifying the fixed terms of the models as regression coefficients are included and a chi-square test based on likelihood can then assess models with differing fixed effects (Hox 2010, 41). The difference between the two estimation-techniques becomes very small for larger groups at the second level (Finch, Bolin, and Kelley 2014, 36).

Because of the differing, theorised effects of constitutional age and comprehensiveness, a test was run to determine the linearity of the relationship between these predictors and the dependent variable. As seen in Figure 5.1, the effect is clearly curvilinear. The effect of age on compliance is similarly curvilinear, just with the opposite trend. As the constitution ages there

seems to be a positive correlation with compliance, until it turns and becomes negative. A figure of this plot can be found in the Appendix. A model specification including squared terms for age and total number of provisions will be included.

Figure 5.1 The effect of comprehensiveness on compliance



5.2.2. Underlying assumptions

There are several assumptions underpinning multilevel analysis. One of which is the violation of independently distributed error terms. This has already been addressed, with the intraclass correlation coefficient indicating a high level of correlation within the error terms, as expected due to the clustering of data within time and within countries. Other assumptions of multilevel modelling usually follow assumptions of linear regression analysis (Hox and Meij 2014, 171). Different assumptions will be addressed after a short discussion on the appropriate number of groups in multilevel models.

There are discussions within the literature concerning the number of groups and observations necessary for the estimates of a multilevel model to be considered as reliable. The recommended number of groups by different scholars can be 10, 30, or 50, or recommendations might not centre on a specific number so much as an understanding of the importance of the

number being “large” (Robson and Pevalin 2016, 27). A low number of groups need not necessitate an abandonment of a multilevel structure, as predicators still exist on different levels and the advantages of the method therefore still holds. What is important is to be aware of potential estimation biases in the level two variables that result from smaller group numbers (Robson and Pevalin 2016, 27; Stegmueller 2013). The number of level one units within the level two groupings are considered to be of less importance (Robson and Pevalin 2016, 27).

Data in the compliance index contains 41 groups if grouped by years, and if grouped by countries contains 175 groups. An empty model with a random effect for countries nested within region was tested as Stegmueller (2013, 753) find evidence that for multilevel models containing large numbers of individual-level observations only, estimates remain robust even when group-level sample sizes are small[14]. This could potentially provide reason to expect that even when grouping countries within regions, given that there are no explanatory variables at that highest level, problems of estimation bias would not severely affect the analysis. As the empty model with random effects for non-nested year and country was preferable, however, there seems no reason to believe the group numbers are too small. Given identical AIC, BIC, deviance, and log likelihood numbers for the model nesting country within region and the model excluding region, there seems to be no additional within-regional variance, that is not already explained by within-country variance.

5.2.2.1 Assumptions of absence of autocorrelation

The typical assumption for multilevel modelling is that the covariance of errors is independent. This assumption should not be assumed to hold for multilevel modelling for longitudinal data (Luke 2004, 70-71). There is already reason to suspect that the error terms might be autocorrelated, due to the clustering of the first level of data within the second level. A Durbin-Watson test is performed, and as it is statistically significant, there indeed seems to be a violation of the assumption of autocorrelation among the residuals.

A way to address the problem of the covariance structure is by changing this structure in the calculation of the models. Only the nlme-package, and not the lme4-package, is capable of handling this however (Finch, Bolin, and Kelley 2014, 96). To address the autocorrelation, an

[14] This finding held for 5 groups at the second level

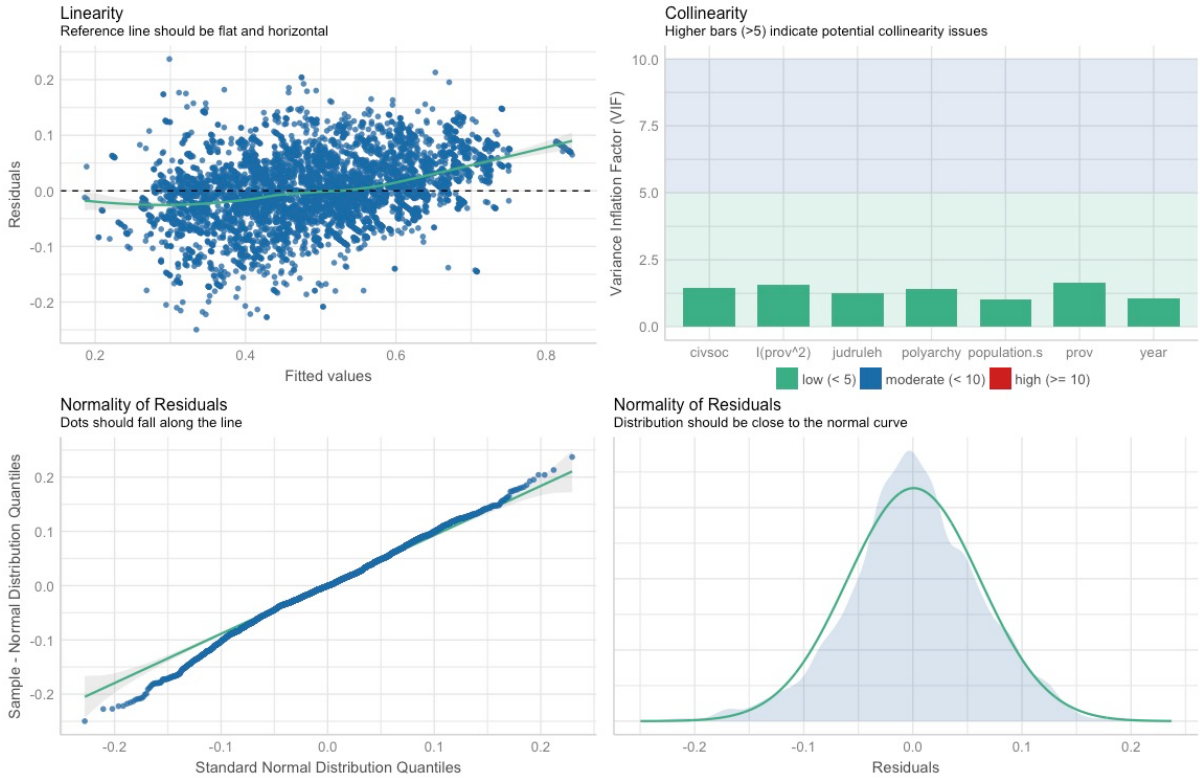
additional model was therefore fitted using lme. The first model was fitted with lme4's lmer function, as that model can fit non-nested models. The second model therefore has no random intercept for year, but as the ICC indicated year might not be a relevant second level group this might provide and interesting contrast to the model which still includes that varying intercept.

Autoregressive error structures are popular when analysing longitudinal multilevel data, as there is an assumed higher degree of correlation for observations closer together in time (Finch, Bolin and Kelley 2014, 96). The autoregressive error structure is both less restrictive than an assumption of a homogenous error structure, and more restrictive than the assumption of unstructured or unrestricted covariance structure. An autoregressive structure is more parsimonious than the unrestricted structure, while not as often a violated assumption as that of a homogenous structure (Hox 2010, 99-103; Luke 2004, 71). As such, the model is fitted with an autoregressive error structure. It seems theoretically reasonable to expect such behaviour from the data; compliance scores for a country are likely very correlated to the compliance score of that country a year previous. This assumption holds true when testing, as the phi coefficient for the correlation structure is 0.91, indicating almost a perfectly correlated relationship.

5.2.2.3 Assumption of homoscedasticity, absence of multicollinearity, and of linear relationships for the residuals

"As with an OLS regression, a mixed-effects model assumes that our residuals are normal and randomly distributed around zero, and that no particular observation exerts undue influence" (Brown 2018, 512). There are many ways of checking for normal distribution of the residuals; the relationship between the actual residuals and theoretic, as inspected in a visual plot, seems relatively alright. The distribution of residuals in the models accounting for the autocorrelation of the error terms are visibly closer to normality than are the two non-nested models. Distribution according to a normal curve show that the two non-nested models have high kurtosis for their residuals, which does not seem to affect the other two models who again are very close to normality. The level two residuals follow the lines more closely.

Figure 5.2 Some visualisations of assumptions for model 3



Visualisations for the other models can be found in the Appendix.

In accounting for the linearity assumption, residuals have been plotted against fitted values, where the expected behaviour is for the residuals to be relatively evenly and randomly placed around the average value of 0 (Hox 2010, 24). Here the relationships seem flipped from the inspection of residual normality. For the two models accounting for autocorrelation, there does seem to be some increased deviation from linearity. The values are slightly more above the line than below for higher fitted values. Overall, the distribution is hopefully not too problematic. The assumption of homoscedasticity for the residuals sees an equal, slight trend away from the line in the higher values for the two nested models. For the two non-nested models the assumption of homogeneity of variance is considered met.

Multicollinearity was a problem due to a lack of centring for the provision-variable. The VIF-test showed it disappeared after a correction of those variables, with the highest value across all four models being 3.05. In the full specification of the model with a random intercept for country there was a collinearity-issue with the variable for the number of provisions present

even after centring, and so that variable was left out of the model. A quadratic term for number of provisions is still left in.

5.3 Results

Table 5.1 Regression-results for all four models

	Non-nested models		**Nested models**	
	Model 1	**Model 2**	**Model 3**	**Model 4**
Constant	.2994*** (.0082)	.3123*** (.0095)	1.3786** (.4399)	1.9756*** (.6989)
Fixed effects				
Number of provisions	-.0026*** (.0004)	-.0013** (.0006)	-.0014*** (.0005)	
Number of provisions squared	.0003*** (.0001)	.0005*** (.0001)	.0001*** (.0001)	.0004*** (0001)
Democracy	.3608*** (.0094)	.3456*** (.0132)	.1873*** (.0095)	.2181*** (.0132)
Population size	-.00545 (.0036)	-.0002 (.0050)	-.0029 (.0062)	.0004 (.0067)
Civil society	.0027 (.0078)	.0009 (.0108)	.0822*** (.0087)	.0749*** (.0112)
Compliance with high court	.0328*** (.0015)	.0356*** (.0022)	.0275*** (.0016)	.0228*** (.0022)
Age		-.0124*** (.0042)		.0011 (.0050)
Age squared		-.0039** (.0017)		-.0078*** (.0022)
High court independence		.0039 (.0024)		.0049* (.0027)
GDP per capita		.0126*** (.0037)		.0090 (.0058)
Civil war		-.0067		.0007

		(.0042)		(.0036)
Year			-.0005**	-.0008** (.0004)
Random effects				
Country	.00851 (.0922)	.00764 (.0874)	.00476 (.0690)	.00432 (.0657)
Year	.00002 (.0047)	.00005 (.0070)		
Residual	.00221 (.0476)	.00225 (.0472)	.00552 (.0743)	.0051 (.0714)
Descriptions				
AIC	-17651.87	-10023.3	-24404.48	-13426.74
BIC	-17585.35	-9931.703	-24331.32	-13335.2
Deviance	-17671.87	-10053.3	-24426.48	-13456.74
N° observations	5724	3316	5724	3316
N° countries	172	152	172	152
N° years	39	27		

Note: *** = p<0.01. ** = p<0.05. * = p<0.1.

Standard errors, or standard deviation for random effects, are in the parentheses.

As seen in Table 5.1 are the results from the regressions for all four models specified. It is fairly evident, from the descriptions of each model, that Model 3 is preferable from a statistical point of view. Models 1 and 3 were the ones fitted based on an exploratory procedure (see for instance Hox 2010, 55-59) where alternately only the fixed and only the random effects were changed and tested. They are the results of several rounds of AIC, BIC and deviance indicating which models were better. Models 2 and 4 were the ones fitted from a purely theoretical point of view. There is more missing data within these two models, as a result of all explanatory variables being present, and thus they have a much smaller number of observations than the other two models. It could be of interest to run such analyses again after imputation and see if the smaller models were still preferred.

The effects of age were excluded from models 1 and 3 during the fittings, but they are included in 2 and 4. Age is significant at a 1 % level and negative in model 2 but positive and non-significant in model 4. It is rather the effect of the quadratic age term that is significant in both models, and the direction of the effect is unchanged. A negative value for a quadratic term means a concave curve, again matching the figure which can be found in the Appendix. This effect seems to strengthen the findings of Elkins, Ginsburg, and Melton (2016) of a maturation effect. It does not lend any strength to theories of decay. Like age, another predictor that changes across models is GDP per capita. It is significant at the 1 % level in model 2, but insignificant in model 4. The effect is positive in both models. Civil society also follows this pattern, but for opposite models. In model 1 and 2 the slight effect is insignificant, whereas in model 3 and 4 it is stronger and significant at the 1 % level.

The predicators of civil war and population size are insignificant across all model specifications. The direction of their effect even changes across the different models. This is contrary to expectations, as civil war and population size are among the "usual suspects" of research on de facto rights respect. It is worth questioning whether this change is due to the ten subcomponents related to structural provisions, or if is something else affecting this outcome. Another unexpected outcome is the effect of year in Models 3 and 4. The effect is significant at a 95 % level, but the direction of the effect is negative. The trend that seemed visible in the bivariate visualisations of Chapter 4 do not apply in these models.

The number of provisions squared, democracy, and compliance with the high court are all significant at the 1 % level for all model specifications. The curvilinear relationship between comprehensiveness and compliance, seen in Figure 5.2, is confirmed by the direction of the coefficient. The higher number of provisions a constitution has, the more the correlation is positive, even if it starts negative. This goes against the findings of Metelska-Szaniawska (2021, 188-189) for post-socialist states, but supports those of Elkins, Ginsburg, and Melton (2016). Compliance with the court mattered much more often than the independence of the court, and the two were negatively correlated at -0.312. This is not quite in line with the expectations from previous research, as independence of the courts typically is an important explainer of respect for rights provisions in practice. The courts' enforcing role do matter to compliance, however, and it could be a question of operationalisation and measures.

That democracy should have a positive and significant effect on constitutional compliance is according to the research on de facto respect for rights provisions and confirms expectations. The measure of electoral democracy, rather than liberal democracy, was used to avoid capturing some accountability structures inherent in the more expansive concept. The effect seems no less significant and while there are reasons for complying with the highest law regardless of regime-type, the result of this analysis indicates there are mechanisms at work in democracies that consistently result in a higher degree of compliance.

While democracy is a significant predictor of constitutional compliance overall, this ought to be accompanied by the caveat that this relationship might be influenced by this rather ubiquitous effect as concerns rights provisions. As we have seen exhibited in the previous chapter, the effect of democracy might be of a less strong or clear nature if testing only for procedural or structural parts of constitutions. It would be of interest to test all the determinants for the different components, separately, given the addition of more subcomponents to the index.

The degree of autocorrelation of the error structures in the data, as well as the very high ICC for country as a varying intercept, indicate that path dependency is a dominant predictor of constitutional compliance. This seems reasonable to expect as an outcome of institutional practices; the average constitution might have what is considered a rather short lifespan, 19 years according to the data of Elkins, Ginsburg, and Melton (2009, 2), but they are still pretty invariant fixtures in our first level of analysis. The degree to which they are complied with is absolutely subject to change, but we don't often expect human rights practices or structural relationships within different branches of government to change drastically each year. This might also explain some of the very low values for the regression coefficients; there is not that much yearly variance that can be affected by anything other than path dependence.

5.3.1 Weaknesses of the model

The testing of assumption showed there were problems with excessive levels of autocorrelation in models 1 and 2. Models 3 and 4 were fitted so as to account for an autoregressive error covariance. The only major change between models 1 and 3, which are then fitted for the same variables except the inclusion of varying intercepts for year in model 1 and year included as a

fixed effect in model 3, were the indicator for civil society engagement. This indicator became significant in the latter model. Other diagnosis of the models showed more kurtosis in the distribution of residuals in models 1 and 2 than in models 3 and 4, and less ideal visualisations for the linearity of models 3 and 4, than for the two former models. Where the two models diverge in their findings ought therefore to be associated with more uncertainty. Additional efforts in imputing might also have made a difference for the more maximalist models 2 and 4; at the very least it could give us a better indication of model fit.

It is important to stress that any findings here are of a preliminary rather than concluding sort. The compliance index is not fully mapped out, especially as it pertains structural provisions, so any results can only reflect the measure as it stands at present. That said, many of the findings are consistent across the different models and the analysis provides at least support to the effect of these predictors. Further analysis into what affects adherence to structural provisions, as well as analysis of compliance as a broader concept than just rights provisions, would be beneficial for knowledge about which conditions foster constitutional respect across differing contexts. Such information could be instrumental in aiding constitution-making processes in the future.

6 Final thoughts

6.1 Discussions about findings

The compliance index gives empirical indications that constitutional compliance takes on different shapes and related differently to surroundings depending on which provisions are assessed for their real-world adherence. This especially concerns the component related to executive powers, which generally inspire much more compliance than the provisions for the judiciary or public administration measured here. While we know more about the respect for rights provisions and state repression, more research into these other parts of the constitution seem necessary.

6.2 Discussions about methodology and potential weaknesses

Any analysis resulting from the current version of this index needs to be interpreted knowing that this is not a complete measure of constitutional compliance. As addressed in Chapter 3, the conceptual validity of the measure suffers from lack of more indicators. It especially suffers from lack of more indicators assessing structural provisions. There is, after all, only a single variable measuring respect for provisions regarding the judiciary. Two provisions measuring the respect for provisions regarding the legislature are also a considerately less than ideal situation. Given that the index contains more rights provisions than structural one, 14 to 10, it is slightly weighted in that direction. The total of ten provisions measured across more structurally oriented provisions, however, provides a more complete measurement of constitutional compliance than one considering only rights provisions, and the measurement might therefore be considered as a step in the direction of measuring all compliance.

My conceptualisation of constitutionalism might not be sufficient. There might be need for additional indicators, providing some sort of qualification as to what can be considered restraining and enabling. The measurement strategy applied in the creating of the index also allows the label of “compliant” to states Sartori (1962) would deem nominal. My stance has throughout this project been simplistic; as long as there are provisions regarding the subject present and they are followed, then there is adherence. And adherence is qualitatively different from “shamness”. This is already explicit in Sartori’s work, however, with the three categories

of sham, nominal and garantiste. The area between merely describing existing power structures and actually creating constraints and structures seem somewhat more elusive, at least in this index. Writing a constitution often is, after all, an exercise in describing where power should and should not reside. If there is considerable attention paid to selecting institutional frameworks that form real limitations on power, perhaps a measure such as this will be enough to capture constitutionalism.

There is also another type of methodology applied in this thesis. In the exploration into some of the likely predictors for the compliance index there were quite strong indications that electoral democracy, compliance on behalf of the government with the high court in decisions it disagrees with, and the convex association with comprehensiveness all affect compliance scores. Further results seemed to indicate that several of the other variables modelled might have an effect, such as the concave association with constitutional age and total number of provisions present in the constitution. The models had some varying results, and due to some weaknesses with them I have chosen to regard them with increased uncertainty. Those results that occurred across all the models, however, should be somewhat more trustworthy. The strongest, and perhaps most unsurprising, finding is perhaps the indication of path dependency in country-year scores for the compliance index. Between-country variation was much greater than that of between-year variation, and the covariance structure had a very high degree of autocorrelation, both suggesting that compliance is related with past compliance.

6.3 Discussions of potential and future applications of a compliance index

As stated in the previous section, the measure as it stands now has its uses, though they are limited by the shortcomings associated with data availability. In my ideal world, there would be expansions and improvements upon the compliance index resulting in a measurement with considerable validity and reliability. Such a measure could be used for insight into how constitutions perform according to its own text in diverse contexts. This has real world applications for recommendations in processes of constitution-making. What institutional arrangements inspire most compliance given certain political, social, and historical conditions? Might certain structures be more unstable in certain conditions? Theories of constitutional functions do suggest, however, that compliance is related to the amount of effort and investment is sunk into the process, and therefore any top-down or outside-in process of

constitution-making might have negative and unintended consequences. Even if they come prepared with all the insights of what typically works elsewhere. Self-enforcement comes in many shapes and are crucial to constitutionalism.

As Elkins, Ginsburg, and Melton (2016, 238) argue, knowing where the gaps are does not give you knowledge of *how* the constitution works. Given that this entire enterprise was motivated out of struggling with definitions of constitutionalism, and that I've operationalised compliance as a mechanism through which constitutionalism works, I'm cannot with any credibility say I believe compliance does not tell us about constitutional functions. Where I do agree is with their argument for assessment of mobilisation around constitutions in hopes of discovering causal mechanisms on when, how, where, and why they do work. Different methods taking aim at clarifying different aspects of how constitutions work can only complement one another. Any quantitative measure on such gaps between de jure and de facto power structures could for instance be useful for locating instances of positive change. And increased knowledge of how constitutions function might further improve measurements of such functions.

6.4 Final thoughts

The aim of this thesis is to engage in a discussion of how to define and measure constitutionalism. It is a concept I struggled with understanding when I first encountered it, as it felt like it kept changing depending on what text I was reading. That frustration turned into fascination, and eventually I had a myriad of new questions that wanted answering. This resulted in an interest in measuring constitutional compliance as a mechanism of constitutionalism, in an attempt at getting closer to measurements of constitutionalism. Such measurements I believe, for both concepts, cannot be without controversy but should not be shied away from on those grounds. The resultant index is lacking, mostly I hope, due to data availability but surely also in some eyes, due to differing opinions on concept and measurement methodology as well as understandings of the underlying concept of constitutionalism. Yet it might give some preliminary insights into what constitutional compliance in a broader sense than respect for rights provisions look like and how it might behave differently.

I have expressed that there might be a need for further expanding on the conceptualisation of constitutionalism, outside of compliance, but would also like to offer up for discussion the

notion that compliance could be sufficient. If the inclusion of more provisions firmly kept focus on what provisions might restrain or enable government power, countries whose constitutions could be classified as nominal in Sartori's framework (1962) would not contain "enough" provisions for achieving the highest scores among compliant countries. The countries with more constitutionalism then comply with more structuring of government power, while those with lesser degrees "merely" comply. Thus, resulting in a lower degree of constitutionalism.

References

Albertus, Michael, and Victor Menaldo. 2013. "The Political Economy of Autocratic Constitutions" in *Constitutions in Authoritarian Regimes* edited by Tom Ginsburg and Alberto Simpser. 53 – 82. Cambridge: Cambridge University Press.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. "Random Effects Structure for Confirmatory Hypothesis Testing: Keep it Maximal." *Journal of Memory and Language* 68 (3): 255-278.

Ben-Bassat, Avi, and Momi Dahan. 2016. "Constitutional Commitment to Social Security and Welfare Policy." *Review of Law and Economics* 12 (1): 165 - 201.

Bolt, Jutta, and Jan Luiten van Zanden. 2020. "Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update" 2020 (10).

Bolt, Jutta, and Jan Luiten van Zanden. 2014. "The Maddison Project: Collaborative research on historical national accounts" *Economic History Review* 67 (3): 627-651.

Brandon, Mark E. 2015. "Constitutionalism." *The Oxford Handbook of the U.S. Constitution*, edited by Mark Tushnet, Mark A. Graber, and Sanford Levinson. DOI: 10.1093/oxfordhb/9780190245757.013.36

Brown, Jonathon D. 2018. "Mixed-Effects Models" *Advanced Statistics for the Behavioral Sciences*. Cham: Springer.

Brown, Nathan J. 2002. *Constitutions in a Non-Constitutional World: Arab Basic Law and Prospects for Accountable Government*. Albany: State University of New York Press.

Chen, Albert H. Y. 2014. "The Achievement of Constitutionalism in Asia." *Constitutionalism in Asia in the Early Twenty-First Century*, edited by Albert H. Y. Chen, 1-31.

Chilton, Adam S., and Mila Versteeg. 2016. “Do Constitutional Rights Make a Difference?” *American Journal of Political Science* 60 (3): 575 – 589.

Christiansen, Jens Peter. 2020. “The Constitution.” The Oxford Handbook of Danish Politics edited by Peter Munk Christiansen, Jørgen Elklit and Peter Nedergaard. S. 8 – 27.

Cingranelli, David, Mikhail Filippov, and Skip Mark. 2021. The CIRIGHTS Dataset. Version 2021.01.21. The Binghamton University Human Right Institute, www.binghamton.edu/institutes/hri/

Cingranelli, David, David L. Richards, and K. Chad Clay. 2014. The CIRI Human Rights Dataset. Version 2014.04.14.

Constitute. 2021. “Constitutions”. https://constituteproject.org/counstitutions?lang=en

Constitute. 2021. “Mexico - Americas - Constitute”. https://constituteproject.org/countries/Americas/Mexico?lang=en

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Anna Lührmann, Seraphine F. Maerz, Kyle L.Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Johannes vonRömer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundtröm, EitanTzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2021. "V-Dem Codebook v11.1" *Varieties of Democracy (V-Dem) Project*.

Davenport, Christian A. 1996. “Constitutional Promises and Repressive Reality: A Cross-National Time-Series Investigation of Why Political and Civil Liberties are Suppressed.” *The Journal of Politics* 58 (3): 627 – 654.

Demidenko, Eugene. 2013. *Mixed Models: Theory and Applications with R*. Hoboken: John Wiley & Sons.

Elkins, Zachary and Tom Ginsburg. 2021. "Characteristics of National Constitutions, Version 3.0." *Comparative Constitutions Project.* Last modified: May 20, 2021. Available at comparativeconstitutionsproject.org.

Elkins, Zachary, Tom Ginsburg, and James Melton. 2016. "Time and constitutional efficacy" *Assessing Constitutional Performance*, edited by Tom Ginsburg and Aziz Huq, 233-267. Cambridge: Cambridge University Press.

Elkins, Zachary, Tom Ginsburg, and James Melton. 2009. *The Endurance of National Constitutions*. Cambridge: Cambridge University Press.

Elkins, Zachary, Tom Ginsburg, and Beth Simmons. 2013. "Getting to Rights: Treaty Ratification, Constitutional Convergence, and Human Rights Practice." *Harvard International Law Journal* 54 (1): 61-96.

Elster, Jon. 2015. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Feld, Lars P., and Stefan Voigt. 2003. "Economic growth and judicial impendence: cross country evidence using a new set of indicators." *European Journal of Political Economy* 19 (3): 497 – 527.

Finch, Holmes W., Jocelyn E. Bolin, and Ken Kelley. 2014. *Multilevel Modeling Using R.* Boca Raton: Chapman & Hall/CRC Press.

Fox, Jonathan, and Deborah Flores. 2009 "Religions, Constitutions, and the State: A Cross-National Study." *The Journal of Politics* 71 (4): 1499 – 1513.

Galligan, Denis J., and Mila Versteeg. 2013. "Theoretical Perspectives on the Social and Political Foundations of Constitutions." *Social and Political Foundations of Constitutions,* edited by Denis J. Galligan and Mila Versteeg, 3-48. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139507509.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Ginsburg, Tom. 2013. "Constitutions as Contract, Constitutions as Charters." *Social and Political Foundations of Constitutions*, edited by Denis J. Galligan and Mila Versteeg, 182-204. Cambridge: Cambridge University Press.

Ginsburg, Tom and Huq. 2016. "Assessing Constitutional Performance." *Assessing Constitutional Performance*, edited by Tom Ginsburg and Aziz Huq, 3-35. Cambridge: Cambridge University Press.

Ginsburg, Tom and James Melton. 2014. "Does De Jure Judicial Independence Really Matter? A Reevaluation of Explanations for Judicial Independence." *Journal of Law and Courts* 2 (2): 187-217.

Ginsburg, Tom and Alberto Simpser. 2013. "Introduction." *Constitutions in Authoritarian Regimes*, edited by Tom Ginsburg and Alberto Simpser, 1-18. Cambridge: Cambridge University Press.

Ginsburg, Tom, James Melton and Zachary Elkins. 2011. "On the evasion of executive term limits." *William and Mary Law Review* 52 (6): 1807 -1872.

Goertz, Gary. 2012. Social Science Concepts: A User's Guide. Princeton: Princeton University Press.

Haber, Stephen and Victor Menaldo. 2011. "Do natural resources fuel authoritarianism? A reappraisal of the resource curse", *American Political Science Review* 105 (1): 1-26.

Hardin, Russell. 1999. *Liberalism, Constitutionalism, and Democracy*. Oxford: Oxford University Press.

Hirschl, Ran. 2013. "The Strategic Foundations of Constitutions." *Social and Political Foundations of Constitutions*, edited by Denis J. Galligan and Mila Versteeg, 151-181. Cambridge: Cambridge University Press.

Hirschl, Ran. 2014. *Comparative Matters: The Renaissance of Comparative Constitutional Law*. Oxford: Oxford University Press.

Holmes, Stephen. 1988. «Precommitment and the paradox of democracy." *Constitutionalism and Democracy*, edited by Jon Elster and Rune Slagstad, 195-240.

Hox, Joop. 2011. "Panel Modeling: Random Coefficients and Covariance Structures." In *Handbook of Advanced Multilevel Analysis*, edited by Joop Hox and J. Kyle Roberts, 85 – 96. New York: Routledge.

Hox, Joop and Leoniek Wijngaards-de Meij. 2014. "The Multilevel Regression Model." In *The SAGE Handbook of Regression Analysis and Causal Inference*.

Keith, Linda Camp, C. Neal Tate, and Steven C. Poe. 2009. "Is the Law a Mere Parchment Barrier to Human Rights Abuse?" *The Journal of Politics* 71 (2): 644 – 660.

King, Jeff. 2013. "Constitutions as Mission Statements." *Social and Political Foundations of Constitutions*, edited by Denis J. Galligan and Mila Versteeg, 73-102. Cambridge: Cambridge University Press.

Law, David and Mila Versteeg. 2013a. "Constitutional Variation among Strains of Authoritarianism." *Constitutions in Authoritarian Regimes*, edited by Tom Ginsburg and Alberto Simpser, 165-196. Cambridge: Cambridge University Press.

Law, David and Mila Versteeg. 2013b. "Sham Constitutions." *California Law Review* 100 (4): 863-952.

Luke, Douglas A. 2004. *Multilevel Modeling*. Los Angeles: SAGE Publications.

Mahoney, James, and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14 (3): 227-249.

McIlwain, Charles Howard. 1947. *Constitutionalism: Ancient and Modern*. Ithaca: Cornell University Press.

Melton, James and Tom Ginsburg. 2014. "Does De Jure Judicial Independence Really Matter?" *Journal of Law and Courts* 2 (2): 187 - 217.

Metelska-Szaniawska, Katarzyna and Jacek Lewkowicz. 2021. "Post-Socialist «Illiberal Democracies»: Do *De Jure* Constitutional Rights Matter?" *Constitutional Political Economy* 32 (2): 233-265.

Metelska-Szaniawska, Katarzyna. 2021. "Post-socialist constitutions: The de jure-de facto gap, it's effects and determinants." *Economics of transition and institutional change* 29 (2): 175-196. https://doi.org/10.1111/ecot.12261.

Møller, Jes Fabricius. 2020. "The Monarch Head of State and National Symbol." The Oxford Handbook of Danish Politics edited by Peter Munk Christiansen, Jørgen Elklit and Peter Nedergaard. S. 46 – 55.

Poe, Steven C., and C. Neal Tate. 1994. "Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis." *The American Political Science Review* 88 (4): 853– 872.

Przeworski. Adam. 2013. "Ruling Against Rules." *Constitutions in Authoritarian Regimes*, edited by Tom Ginsburg and Alberto Simpser, 21-35.

Robson, Karen and David Pevalin. 2016. *Multilevel Modeling in plain language*. London: SAGE Publications.

Sartori, Giovanni. 1962. "Constitutionalism: A Preliminary Discussion." *The American Political Science Review* 56 (4): 853-864. https://doi.org/10.2307/1952788.

Stegmueller. Daniel. 2013. "How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches." *American Journal of Political Science* 57 (3): 748-761.

Stortinget. 2020. "Om Grunnloven." Oppdatert 14.05.2020. https://www.stortinget.no/no/Stortinget-og-demokratiet/Grunnloven/om_grunnloven

Strauss, David A. 1996. "Common Law Constitutional Interpretation." *The University of Chicago Law Review* 63 (3): 877 – 935.

Sunstein, Cass R. 1988. "Constitutions and democracies: an epilogue." *Constitutionalism and Democracy*, edited by Jon Elster and Rune Slagstad, 327-356.

Thio, Li-Ann. 2012. "Constitutionalism in Illiberal Polities" *The Oxford Handbook of Comparative Constitutional Law*, edited by Michel Rosenfeld and András Sajó, 133-154.

Tushnet, Mark. 2014. *Advanced Introduction to Comparative Constitutional Law*. Edward Elgar: Cheltenham.

Tushnet, Mark. 2013a. "Authoritarian Constitutionalism" *Constitutions in Authoritarian Regimes,* edited by Tom Ginsburg and Alberto Simpser, 36-50. Cambridge: Cambridge University Press.

Tushnet. Mark. 2013b. "Constitution-Making: An Introduction.» *Texas Law Review* 91 (7): 1983-2013.

Versteeg, Mila, Timothy Horley, Anne Meng, Mauricio Guim and Marilyn Guirguis. 2020. "The Law and Politics of Presidential Term Limit Evasion". *Columbia Law Review* 120 (1): 173 - 248.

Voigt, Stefan, Jerg Gutmann, and Lars P. Feld. 2015. «Economic growth and judicial independence, a dozen years on: Cross-country evidence using an updated Set of indicators." *European Journal of Political Economy* 38 (38): 197 – 211.

West, Brady T., Kathleen B. Welch and Andrzej T. Galecki. 2014. *Linear Mixed Models*. New York: Chapman and Hall/CRC.

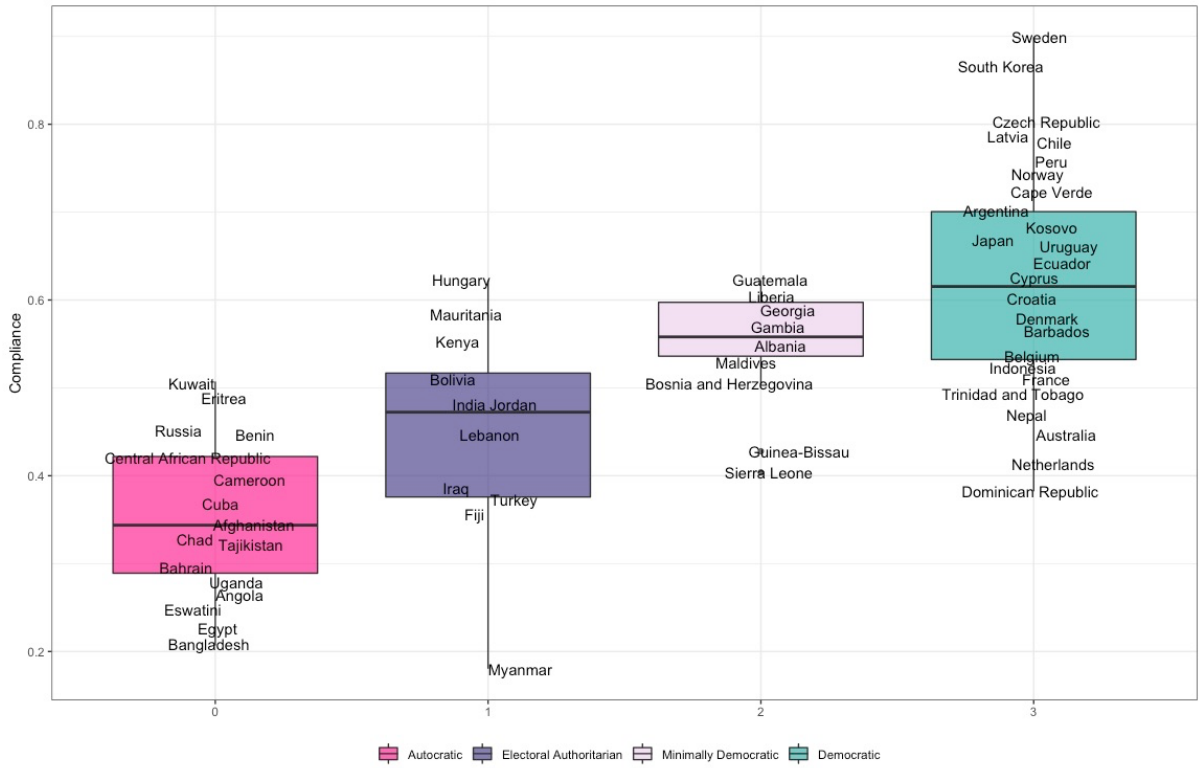
Western, Bruce. 1998. “Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Aproach.” *American Journal of Political Science* 42 (4): 1233 – 1259. URL: https://www.jstor.org/stable/2991856

World Bank. 2019. *World Development Indicators 2019*, World Bank, Washington.

Appendices

Appendix A

Figure 1 Compliance scores by the V-Dem Electoral Democracy Index



Appendix B

Figure 1 Visualisations for regression assumptions for model 1

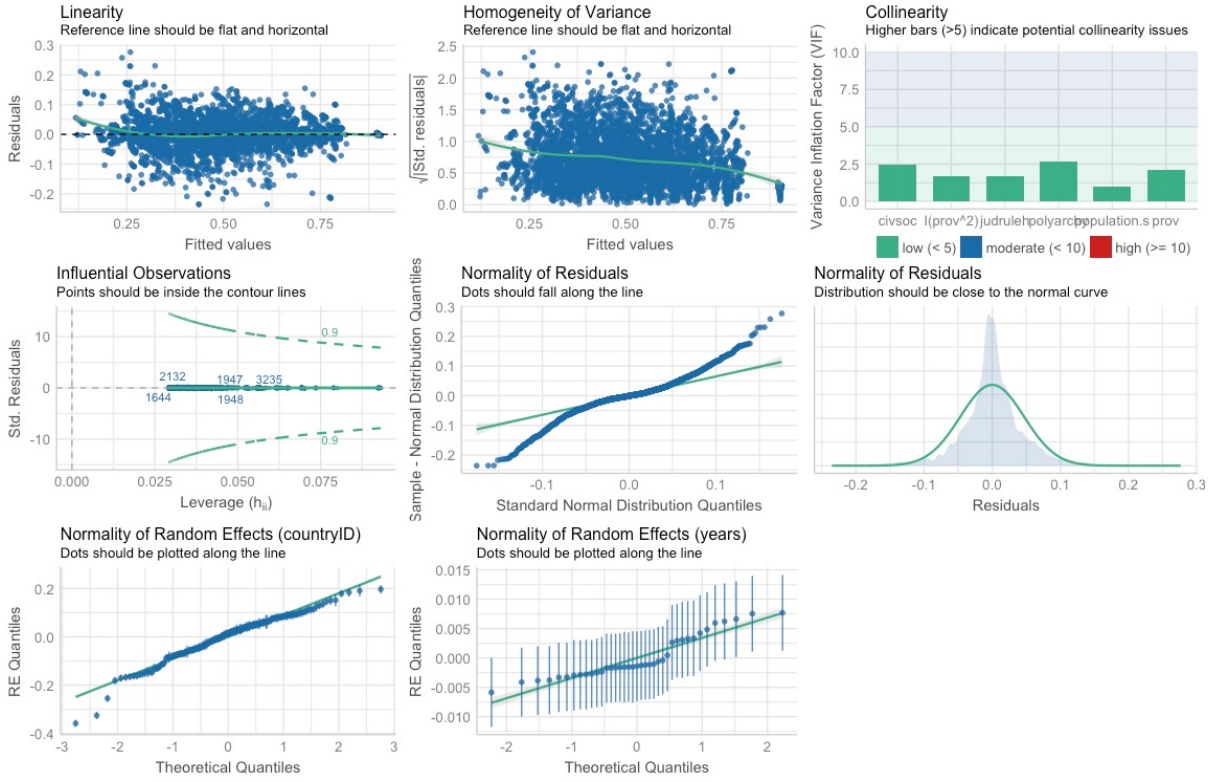


Figure 2 Visualisations for regression assumptions for model 2

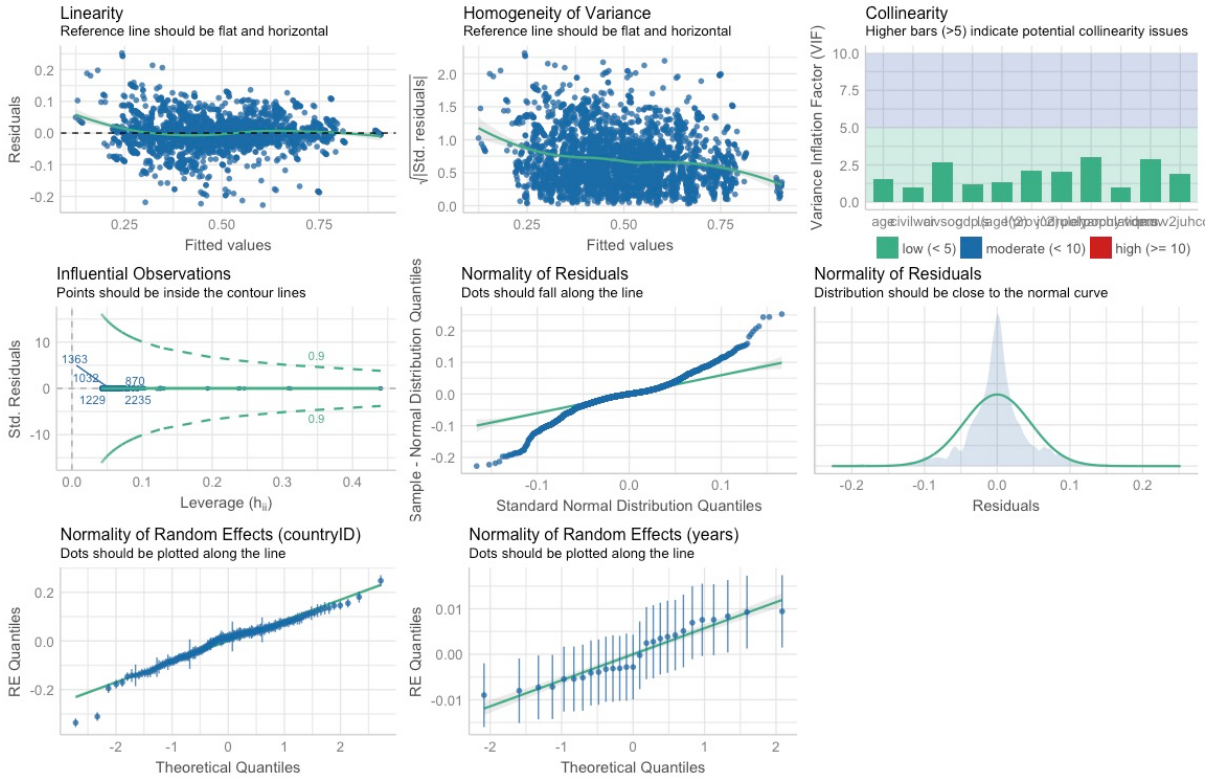


Figure 3 Visualisations for regression assumptions for model 4