University of Oslo

Master of Philosophy in Peace and Conflict Studies

The Same, Only Different:

Revisiting State Capacity's Theoretical and Empirical Implications for the Study of Intrastate Armed Conflict

Charlotte Tandberg

Supervised by Professor Scott Gates & Professor Indra de Soysa

Spring 2021

The Same, Only Different:

Revisiting State Capacity's Theoretical and Empirical Implications for the Study of Intrastate Armed Conflict

Charlotte Tandberg

Peace and Conflict Studies
Department of Political Science
Faculty of Social Sciences
University of Oslo

* * *

Spring 2021

The Same, Only Different: Revisiting State Capacity's Theoretical and Empirical Implications for the Study of Intrastate Armed Conflict

http://www.duo.uio.no/

Print: Reprosentralen, University of Oslo

Word count: 33,053

Abstract

The extent to which state capacity serves as an explanation for variations in intrastate (in)stability has for long been a major focus of political science in general, and armed conflict research in particular. Despite the conceptual and theoretical richness of the state capacity literature, this thesis argues that its measurements suffer from significant flaws. By reconceptualizing state capacity based on three identified dimensions and drawing upon 20 carefully selected indicators, exploratory factor analysis is employed with the aim to discover whether indicators form regular patterns and vary together. The results provide strong support for the proposition that the theorized dimensions are interrelated and hard to distinguish empirically.

From there, a new aggregate measure of state capacity is constructed with demonstrated validity and utility for use as a powerful predictor of intrastate conflict onset. Founded on a solid theoretical base in which construct validity is critically assessed and capturing only the core state functions that are minimally necessary for the modern state to implement desired policies, the measure should capture the underlying concept more comprehensively than previous works. Thus, this thesis represents one of the first steps towards bridging an important methodological gap in the literature and provides the basis for a measurement strategy that can be applied not only to studies on peace and conflict, but to all comparative cross-national research concerned with the causes and consequences of state capacity.

Acknowledgements

With the submission of this thesis, my two years in PECOS are drawing to a close. Despite a turbulent year and highly unforeseeable challenges caused by the COVID-19 pandemic, there are a number of people who have contributed to making the dissertation-writing process enjoyable. They all deserve a special thanks. First of all, I want to thank my two supervisors, Scott Gates and Indra de Soysa, for invaluable guidance, encouragement, and interesting discussions. Thank you, Gates, for being inspiring from the very first moment. Thank you, de Soysa, for taking interest in the project, for participating with your knowledge on short notice, and for always responding quickly to all frustrating emails, even on Saturday nights. I am extremely thankful for working with both of you. I would also like to thank Jo Jakobsen for the encouragement and for recommending de Soysa as a second supervisor.

Furthermore, a special thanks goes to all those who have offered unconditional support in other ways. Emil Hillesøy, thank you for sharing your brilliant Excel- and R-skills, and for being there throughout it all. To all fellow PECOS students, this process would not have been the same without you. I appreciate all coffee breaks, UNO games, internal jokes, and the close friendships we have developed along the way. Being part of this community has been truly rewarding. Lastly, I want to thank my mom and grandpa for always being supportive, for stressing the importance of dedication, and for being only a phone call away when needed.

All remaining errors in this thesis are my responsibilities alone.

Charlotte Tandberg

Oslo, Spring 2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The extent to which state capacity serves as an explanation for variations in intrastate (in)stability has for long been a major focus of political science in general, and armed conflict research in particular. For instance, researchers have repeatedly investigated how national attributes like economic development, regime type, extractive and redistributive efforts, bureaucratic quality and surveillance control operate in affect propensity for armed conflict (see e.g. Hegre, Ellingsen, Gates, & Gleditsch, 2001; Fearon & Laitin, 2003; P. Collier & Hoeffler, 2004; Fjelde & de Soysa, 2009; Besley & Persson, 2009; Sobek, 2010; Cederman, Gleditsch, & Buhaug, 2013; Koren & Sarbahi, 2017; Mason & Greig, 2017; Bakaki, 2020). In fact, the majority of such studies conclude that states' abilities to mitigate favorable conditions for insurgency, either by repressive or accommodative means, are pivotal determinants of civil conflict likelihood and dynamics.

Although these findings are important in their own rights, this thesis argues that, despite the conceptual and theoretical richness of the state capacity literature, its measurements suffer from significant flaws. It further suggests that too little attention is paid to whether the current spectrum of measures even capture the same underlying concept. Conflicting conceptualizations failing to take construct validity into account have led a significant number of state capacity measures to either be aggregated into often questionable catch-all indices, or to be be broken down to highly diaggregated sets of state functions, distracting attention from the

state's most fundamental *capacities* and ignoring the interplay between different dimensions of the state. Thus, a core question confronting armed conflict researchers is how to empirically address the concept's latent nature and multidimensionality (Soifer, 2008; Hendrix, 2010; Lindvall & Teorell, 2016; Hanson & Sigman, 2020).

In this thesis, I attempt to address the existing methodological inconsistence theoretically and empirically. Drawing on the burgeoning state capacity-civil conflict literature, I review competing definitions and operationalizations, assess their construct validity, and delineate three dimensions of capabilities that states should hold in order to fulfill their most fundamental functions and thus avoid conflict onset. These dimensions - which I argue are deeply interrelated, codependent, and hard to distinguish empirically - exist within the state's organizational structures independently of political decisions on how/whether to prioritize them, and account for coercive, bureaucratic-administrative, and extractive capacity.

Taking as point of departure a thorough reconceptualization, I thereafter explore the presumed interrelationship between 20 carefully selected indicators relating to state capacity's three dimensions by employing principal factor analysis. The data covers 165 countries for up to 29 years (1991-2019), resulting in 4,722 country-year observations. Results provide strong support for the proposition that the theorized dimensions are interrelated also in practice: low underlying dimensionality leads one factor to explain close to 67% of the variance in the measures, and to a large extent, this factor captures the bundle of qualities that were hypothesized to make up the state's abilities to carry out its core functions. Finally, the factor is transformed into a new, aggregate state capacity measure with demonstrated validity and utility as a strong and consistent predictor of intrastate conflict outbreak. The remainder of this first chapter introduces the building blocks on which this thesis is built, and elaborates on the research question it is guided by:

> *"To what extent do existing operationalizations of state capacity accurately and independently capture the theoretical concept, and how may instead an aggregate measure of core state functions contribute to our understanding?"*

1.1 The Role of State Capacity

According to the Uppsala Conflict Data Program, the number of state-based armed conflicts in 2019 was 54. Of these, only two occured between states, making intrastate armed conflicts the primary form of organized violence in the contemporary world (N.-P. Gleditsch, Wallensteen, Eriksson, Sollenberg, & Strand, 2002; Pettersson & Öberg, 2020). On the peculiarities of the intrastate genre, Fearon and Laitin (2003) have emphasized that civil wars in the post-World War II era were far more deadly, harder to resolve and long-lived than their interstate counterparts. Furthermore, and rather not surprisingly, Gates, Nygård, and Strand (2012) found strong negative effects of intrastate conflict on central development issues such as poverty, education, access to nourishment, and health care.

Additionally, P. Collier et al. (2003) illustrate that many of the economic, political and social costs of war continue to accrue long after the initial conflict has ended, preventing states from rebuilding their different apparatus. This has the potential to leave states stuck in a "conflict trap" of reversed development, making the state more prone to further conflict (ibid.). Moreover, transnational dimensions of within-state conflicts include neighborhood contagion (Buhaug & Gleditsch, 2008; Braithwaite, 2010), refugee movements (Salehyan & Gleditsch, 2006) and ideological radicalization (Schubiger & Zelina, 2017; Walter, 2017). The fact that civil conflicts cluster geographically also suggests that conflicts are not independent of each other, which is a pressing challenge with implications for international stability, security, and interventions (K. Gleditsch, Dorussen, Metternich, & Ruggeri, 2012).

Consequently, academic interest in understanding and explaining intrastate armed conflicts – and particularly, what causes them – has increased tremendously. Out of the 82 countries that had a state-based conflict with $\geq$ 25 casualties between 1989 and 2019, 40% are African, and an additional 22% are placed on the belt stretching form Syria to the Philippines[1] (Sundberg & Melander, 2013). Since 2010, 38 states have experienced such conflicts, and 86% of them are located either

[1]Data is based on the UCDP Georeferenced Event Dataset (GED) Global version 20.1, and can be downloaded here `https://ucdp.uu.se/downloads/`

in Africa or on the same Asian belt. The common attribute for these regions is that states are often recognized for their governments' inability to control territory, regulate violence, and to implement (conflict-reducing) policies. This has encouraged a burgeoning body of research on the so-called state capacity-armed conflict nexus, privileging structural factors to explain variation in intrastate (in)stability (Hendrix, 2010). In general terms, these factors are akin in the way that they seek the causes of civil conflict in the 'conditions that favor insurgency' (Fearon & Laitin, 2003), or the structural environments that enable rebel groups to organize, fight and exist.

The state capacity model emphasizes the role of the state in mitigating these conditions, normally in four aspects - of which none are easily distinguishable theoretically nor empirically. The first relates to the state's ability to physically access and project authority throughout its territory (e.g. Ricciuti & Rossignoli, 2017); the second emphasizes the state's coercive capabilities by way of its police and military (e.g. Walter, 2006); third, the state's level of bureaucratic and administrative sophistication and ability to implement policies (e.g. Goodwin & Skocpol, 1989); and fourth, its ability to extract the resources and revenue necessary to accomplish their desired goals (e.g. Thies, 2015). Not surprisingly, the majority of studies find significant evidence that stronger states (manifested in high state capacity) have a decreased risk of experiencing civil conflicts (see e.g. Sobek, 2010; Newman, 2014).

1.2 Defining State Capacity

As point of departure in finding a satisfactory definition of state capacity, I have chosen to avoid conceptualizations that are concerned with either normative standards on how state power is exercised or societal causes and outcomes that are influenced by events exogenous to the state's immediate control - though, of course, they are closely related. Nevertheless, such aspects are conceptually distinct from capabilities "that exist within a state's organizational structures, and the territorial reach of these capacities" (Hanson & Sigman, 2020, p.3).

Instead, it has been helpful to unpack Mann's (1984, p.189) concept of the state's infrastructural power: the capacity of the state to penetrate society and to

"implement logistically political decisions throughout the realm." Similarly, Skocpol (1985, p.16) contends that "sheer sovereign integrity and the stable administrative-military control of a given territory are preconditions for any state's ability to implement policies" and that "loyal and skilled officials and plentiful financial resources are basic to state effectiveness in attaining all sorts of goals". These definitions are beneficial in a number of ways, not least because they acknowledge the importance of the state's *power* in order to 'get things done' independently of which sets of decisions or outcomes they relate to. Furthermore, they combine power with both physical, human, and administrative abilities that all state outputs ultimately depend on (Dahl, 1957; Migdahl, 1988; Geddes, 1996; Lindvall & Teorell, 2016).

Hence, the 'general underpinnings' which I consider pivotal for the state to fulfill its most fundamental functions relate to 1) its coercive abilities to deter and repel challengers with force, to maintain internal order and to ensure compliance with the law, 2) the bureaucratic-administrative sophistication necessary to collect, monitor and implement data, information and policy, to reach citizens and to conduct decision-making in an impartial manner, and 3) the extraction of revenue, in order to support all other activities (Hegre, 2018; Hanson & Sigman, 2020; Gjerløw, Knutsen, Wig, & Wilson, 2021). Although conceptually distinct, they are interrelated, and significant strength in one should require at least some strength in the others. High state capacity in this thesis therefore equates with the ability to effectively implement whatever policies policymakers decide on with reference to the three dimensions outlined above, and at the lower end of the scale, one should expect to find states already mired in conflict and/or lacking what is referred to as "stateness" (Linz & Stepan, 1996).

1.3 Prospects of Contribution

There is nothing new to the fact that researchers disagree over the relative importance of the above-mentioned dimensions, but implications arise when different approaches argue that their one-dimensional measure of state capacity works as a strong one-size-fits-all proxy, ignoring first of all the multitude of core state functions

that does not relate to their particular inquiry, and second of all, other potential causal mechanisms associated with the measure[2]. In this way, the lacking consensus on how to measure e.g. state capacity within equal, or even the same, theoretical definitions rises important questions relating to essentially all aspects of validity and thus research quality and trustworthiness (Carmines & Zeller, 1979; Adcock & Collier, 2001). For this reason, Blalock (1979, p.882) has noted that "the most serious and important problems that require our immediate [...] attention are those of conceptualization and measurement."

Thus, revisiting the state of state capacity's conceptualizations, operationalizations and construct validity is an important task per se. By estimating state capacity as a latent variable only capturing the very core state functions that are minimally necessary to implement desired policies, I hope to 1) reconceptualize state capacity and illustrate the interrelationship between the three dimensions, and 2) demonstrate why the new aggregate measure is a better alternative for conflict analyses in terms of validity and utility than previous attempts, often conflating states' material resources and organizational competencies with conceptually distinct features. As such, the main motivation for this thesis is to contribute to the field with a potentially more valid and comprehensively conceptualized measure, offering a basis for a measurement strategy that can be applied not only to studies on peace and conflict, but to all comparative cross-national research concerned with the causes and consequences of state capacity.

[2]GDP per capita is one such example. For instance, Fearon and Laitin (2003, p.80) contend that "a higher per capita income should be associated with a lower risk of civil war onset because it is a proxy for a state's overall financial, administrative, police, and military capabilities." However, Fjelde and de Soysa (2009), Hendrix (2010) and Jakobsen, de Soysa, and Jakobsen (2013) stress the fact that GDP per capita also can proxy economic grievances and the opportunity cost to participating in violence. The ambiguity in interpretation thus suggests that the measure is a questionable proxy for state capacity because it captures other plausible causal mechanisms.

1.4 Methods and Findings

The most central question in this thesis is whether existing, disaggregated operationalizations of state capacity accurately and independently capture the theoretical concept, or if an aggregate measure of clustered characteristics that relate to core state functions is empirically preferable, particularly in relation to intrastate conflict onset. This question is attempted answered in three steps. First, I assess prevailing conceptualizations and operationalizations on the basis of construct validity, and discuss the bigger implications when attempting to measure a latent variable such as state capacity. I argue that current applications of state capacity in quantitative analyses suffer from significant flaws both theoretically and empirically. Arguing that especially three dimensions are interrelated, codependent, and hard to separate in practice, I reconceptualize the state capacity concept by taking as point of departure the state's material resources and organizational competencies that exist independently on political decisions on how to deploy them.

Second, I employ data from 20 carefully selected indicators in an exploratory principal factor analysis. The indicators are drawn from multiple sources and span 29 years (1991-2019) for up to 165 countries in a given year[3]. To compensate for missing data, I undertake predictive mean matching by multiple imputation. The factor analysis yields results confirming that the dimensionality of the presumed underlying concept is low, and rejects the hypothesis that coercive, bureaucratic/administrative, and extractive capacity can be empirically distinguished. The obtained factor of interest is termed State Capacity, and its country rank ordering favors bureaucratic, technologically advanced countries that are capable of ensuring compliance with the law normally by other means than the coercive, and extracting information and taxes that require higher levels of societal penetration. The only factor to represent a distinct state capacity dimension was factor 3, on which indicators representing devotion to the military and natural resource dependence load heavily. This factor is positively associated with conflict onset.

[3]An overview of data coverage and a qualitative description of all indicators can be found in Appendix Table A.1 and Appendix Figure A.1, respectively.

Lastly, I attempt to validate the new measure by running different types of validity and utility checks. It is examined in terms of its face validity, content validity, convergent validity and predictive validity. The findings provide evidence of high validity and demonstrated utility of the variable as a strong predictor of intrastate conflict outbreak. The latter is tested by first replicating Fearon and Laitin (2003), and second by running a logistic regression controlling for a variety of other aspects to the state that generally are considered decisive for whether or not a country will experience conflict. Thus, the most central result is a new, thoroughly conceptualized aggregate measure with broad utility that should be relevant for multiple research traditions beyond the peace and conflict sphere. Because research only can be relevant and reliable if approaches to modeling are enhanced, this thesis represents an important step towards bridging the current gap by addressing key conceptual and measurement issues in the state capacity-armed conflict nexus.

1.5 Roadmap of the Thesis

This thesis contains seven chapters. Chapter 2 presents the basis for the theoretical agenda, and provides an overview of current trends in armed conflict and their explanations, underscoring why state capacity is of importance in mitigating favorable conditions for insurgency. Chapter 3 builds on this argument, and emphasizes that despite the general agreement that state capacities are critical to maintaining civil peace, the lack of consensus on how to define and measure the concept rises important questions relating to validity and thus research quality and trustworthiness. After reconceptualizing state capacity, the chapter identifies 20 distinct indicators and assesses each measure according to two aspects of construct validity, theoretically arguing in favor of an aggregate measure. Chapter 4 develops a quantitative research design and elaborates the methods used to statistically test whether the indicators form regular patterns and vary together. Here, the data is introduced, methods outlined, methodological choices explained, and implications addressed. In chapter 5, the results of the factor analysis are presented, and the new state capacity measure is announced. This measure is examined more thoroughly in chapter 6,

which attempts to fully validate it. The chapter also demonstrates the utility of state capacity as an aggregate measure using the factor-estimated levels of state capacity as predictors of intrastate conflict outbreak. Finally, chapter 7 summarizes the thesis' findings, contributions, and limitations.

Chapter 2

Background

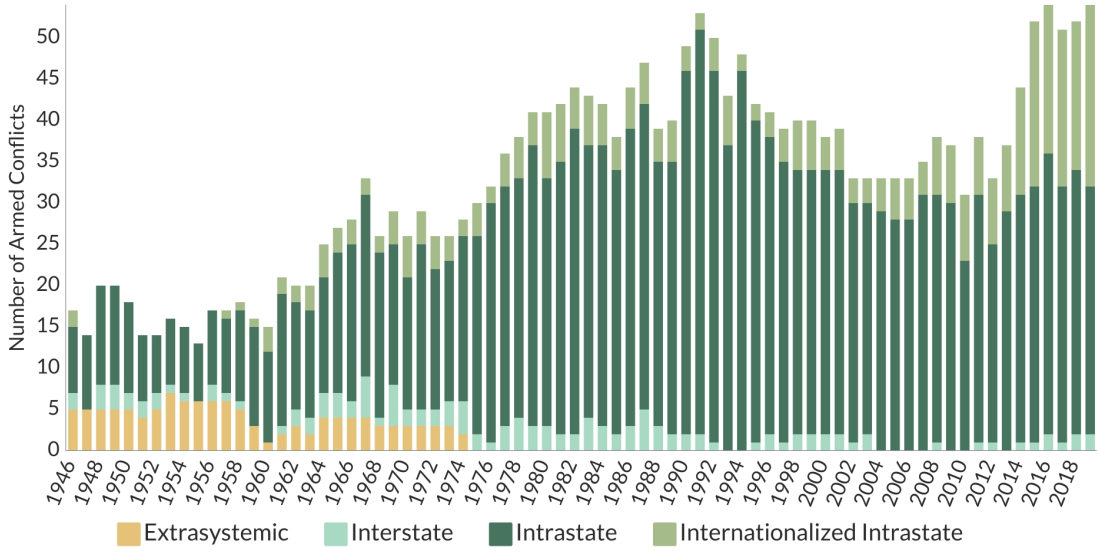
What determine the dynamics of violent conflict between government forces and armed opposition movements? Where are these conflicts situated, and which states can be said to be at greater risk? A considerable amount of researchers have since the 1960s systematically investigated cross-national variations in intrastate (in)stability, and many have offered influential insights. Particularly three explanatory logics bore fruit to the theoretical agenda, of which one has become the dominant: the state capacity model, focusing on the structural environments that enable dissident groups to organize, fight, and exist. This chapter first presents the current trends in conflict and reviews the leading explanatory literature. Thereafter, it scrutinizes how approaches to state building and the "stateness-first" argument have formed our understanding of political order in general and state capacity in particular.

2.1 Trends in Conflict and its Explanations

Armed conflicts between states have historically accounted for some of the deadliest affairs in human history. Since the end of the Second World War however, interstate wars have become rare events. Instead, internal conflicts have become much more common. In the second half of the 20^{th} century, the number of intrastate armed conflicts by far surpassed their interstate counterparts, whilst also becoming increasingly internationalized (N.-P. Gleditsch et al., 2002). Furthermore, while

casualty figures for intrastate conflicts have decreased steadily since 2014, mainly because of a reduction in violence in Syria, the number of active conflicts has continued to rise and the number of conflict locations has remained high and consistent. These trends are illustrated in Figure 2.1 and Figure 2.2.

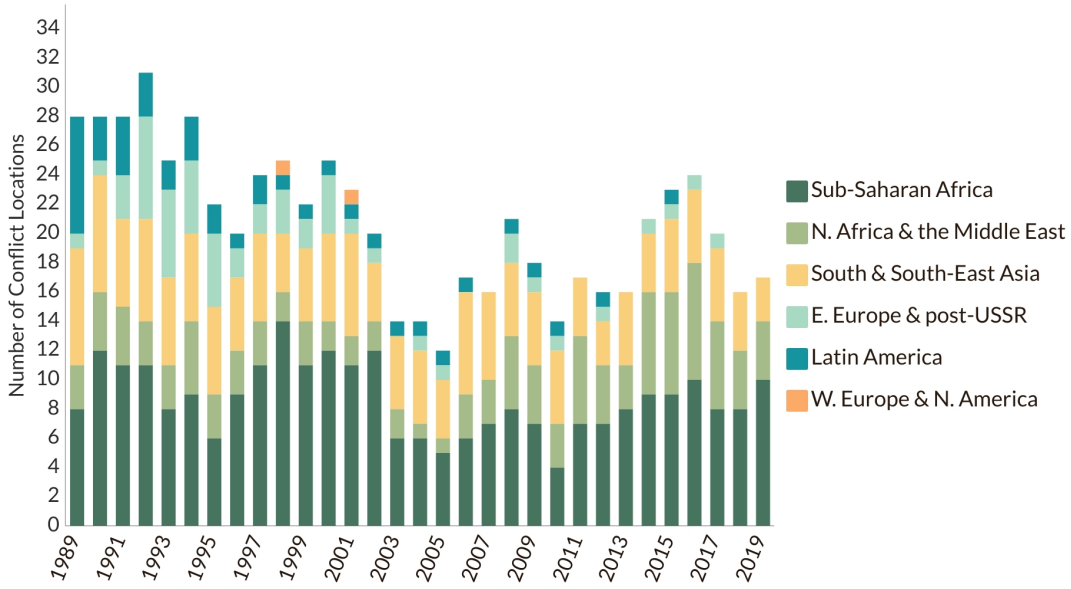
FIGURE 2.1: Number of armed conflicts by type, 1946-2019



Source: UCDP/PRIO Armed Conflict Dataset v20.1 (N.-P. Gleditsch et al., 2002; Pettersson & Öberg, 2020)

According to UCDP and Figure 2.1, the number of state-based, intrastate armed conflicts in the world in 2019 was 52, matching the post-Cold War peak in 2016 (N.-P. Gleditsch et al., 2002; Strand, Rustad, Nygård, & Hegre, 2020). A closer look at the numbers suggests that these conflicts cluster geographically, and that the occurrence of one is unlikely to be independent of another. As can be seen from Figure 2.2, 17 states had one (or more) state-based conflict with ≥ 25 casualities in 2019, of which 11 are African. Out of the 82 countries that have had suchlike conflicts between 1989 and 2019, 40% are African, and an additional 22% are placed on the belt stretching form Syria to the Philippines (Sundberg & Melander, 2013). Since 2010, 86% of conflict locations have been either African, Middle Eastern, or in-between Afghanistan and the Philippines.

FIGURE 2.2: Number of conflict locations by geopolitical region, 1989-2019



Source: UCDP Georeferenced Event Dataset Global v20.1 (Sundberg & Melander, 2013; Högbladh, 2020)

The majority of conflict onsets since the mid-1990s have furthermore not been characterized by new, but recurring conflicts. The disturbing story told by Gates, Nygård, and Trappeniers (2016) reveals that around 60% of conflicts recur, and that the two regions most notably contributing to the recurrence-trend are the Middle East and North Africa (MENA) and Sub-Saharan Africa regions. Without strippingthese regions of neither history nor context, it should be safe to note that the majority of states in which conflicts regularly reoccur are disproportionately poorer and recognized by their governments' lacking incentives and/or capabilities to implement policies and regulate violence (Hegre, 2018). Unsurprisingly, armed conflicts also tend to exacerbate the conditions that helped cause them in the first place: war wrecks the economy, decreases state capacity, and increases the risk of further war (Von Einsiedel, Bosetti, Cockayne, Salih, & Wan, 2017; P. Collier et al., 2003).

2.1.1 The Holy Trinity: Grievances, Greed, and Opportunities

It is generally understood that civil conflicts arise from a variety of factors. Despite this, an overwhelming number of quantitative studies employ a constricting explanatory logic that classifies these causal factors into three categories, namely grievances, greed, and opportunities (Sobek, 2010, p.268; Ward, Greenhill, & Bakke, 2010, p.364; Cederman & Vogt, 2017, p.1995). The two former are both 'rebel-centric' approaches relating to the willingness of actors to utilize violence in order to alter power relations, while the latter presupposes certain structural environments providing potential rebels with opportunities to revolt.

The point of departure for explanations centering grievances stem from the influential work of Gurr (1970) and his relative deprivation theory, focusing on state-imposed disadvantages and group discrimination. The theory views collective violence as a reaction to frustration stemming from the failure to achieve aspired goals (see also Davies, 1962; Horowitz, 1985). However, because of its mixed records in terms of empirical evidence, the frustration-aggression theory has attracted criticisms from among others Skocpol (1979), Fearon and Laitin (2003), P. Collier and Hoeffler (2004) and Tilly (1978), arguing that grievances are ubiquitous and therefore an inadequate explanation of rare events like civil wars. As a consequence, scholars like Stewart (2008) and Cederman et al. (2013) have reversed the criticism, contending that such conclusions are premature. Østby (2008) makes the case that interpersonal measures dismisses the fundamental importance of 'horizontal', or group-level, identity-based inequalities - which she finds is positively related to conflict outbreak. Accordingly, it is not vertical inequalities, ethnicity nor ethnic heterogeneity per se that makes a state more prone to intrastate conflicts, but instead distributional group-asymmetries related to socioeconomic advantages and political opportunities (see also Wucherpfennig, Metternich, Cederman, & Gleditsch, 2012; Hillesund et al., 2018).

On the other hand, greed-based accounts are concerned with the economics of conflict, and view internal conflict as a consequence of individuals' evaluation of the

costs and benefits of organizing armed violence on the one and and their desire to maximize their profits, normally in a material sense on the other (C. Anderson & Carter, 2019; de Soysa, Vadlamannati, & Finseraas, 2021). For instance, P. Collier and Hoeffler (2004) find strong empirical support for the proposition that the availability of natural resources motivates rapacious behavior. de Soysa (2000) adds that an abundance of natural resources is seen to act as a 'honey pot' creating incentives for capture and control, thus increasing the propensity for causing armed conflict. Indeed, as exemplified by e.g. de Soysa et al. (2021), many of the post-Cold War conflicts in Sub-Saharan Africa seemed to be driven by "loot-seeking" rather than "justice-seeking" motives. The wars in Liberia, Sierra Leone, Angola and Colombia are commonly propounded as arising from or fueled by revenue-generating opportunities from controlling resources like diamond mines, oil, timber, and drugs. Rebel leaders and other dissidents are according to this logic assumed to calculate the cost and benefits of resorting to insurrection, and contend that one should expect civil conflicts where populations have large 'youth bulges', education is limited, opportunity costs are lower and where probabilities for rebel financing and personal enrichment is higher (see also Ross, 2015; Dal Bó & Dal Bó, 2011).

Whether the cause of conflict is attributed to greed or grievance, their shared characteristic is the perception of a particular deprivation. This sense of deprivation can be leaning towards the more individualist and economic on the one hand, or towards group-based identities on the other - but either way, the deprivation is both caused and sustained by an interplay between the two. Keen (2008a, 2008b) demonstrates this interaction using examples from Sudan and Sierra Leone, where grievances were 'useful' to governments in various ways. The civil war in Sri Lanka too is a useful example, where the Tamils' perceived grievances during Sinhalese rule later were interwoven with greed factors like entitlements to resources.

The last aspect of the Trinity doctrine seeks the causes of civil conflict not so much in rebels' motives, but in the structural environments that enable rebel groups to organize, fight and exist. These factors tend to clump together, but in the wider literature, rebels' opportunities to revolt by large depend on the government's capacity to penetrate society, control territory, and effectively implement political

decisions throughout the realm. Perhaps the most influential explanation within the opportunity-approach is proposed by Fearon and Laitin (2003, p.75), stressing that insurgent violence is more likely to erupt in weaker states because these states foster 'favorable conditions for insurgency' (see also e.g. Tilly, 1985).

Although originally applied to explain the eruption of conflict in the international system, Organski and Kugler (1980) present the power-transition theory, which insists that the source of war is to be found in the differential rate of growth of the two most powerful nations in the system: the dominant and the challenger. If one controls for size of nation states, "the rest of the differences in power can be accounted for by differences in levels of development [...] and most important are economic productivity and the efficiency of the political system in extracting and aggregating human and material resources into pools available for national purposes" (1980, p.21; see also Kugler & Lemke, 1996). At the heart of the theory lies the assumption that the major source of power for a state is its own sosioeconomic and political development, and predictions of who will win depend not only on the power potential of a nation, but also on the state's capacities to *mobilize* its resources.

Furthermore, because "weak states" are recognized by among other things limited reach of state institution and lack of security provision, theories on the security dilemma can be useful in explaining why emerging anarchical environments in the absence of central authority offer breeding grounds for outlaw groups (Herz, 1951; Jervis, 1978; C. Kaufmann, 1996). Although intentionally developed with interstate relations in mind, the baseline theory can well be applied to the intrastate genre. Where central authority is lacking, citizens are forced to provide for their own security, which enables potential rebels to 'organize in the shadows' and to acquire sufficient attributes of sovereign states. In turn, this threatens both the authority of the central state and the security of other groups in society, thus increasing the risk of the security dilemma and unconstrained violence[4].

[4]There are multiple theories normally applied to interstate relations that also could be of relevance here (e.g. the theories of Starr, 1978; Fearon, 1995; Arreguín-Toft, 2005; Jervis, 1976; G. Snyder, 1971; Levy, 1987). However, they will not be dealt with in greater detail, as they are normally not directly concerned with the role of state *capacities*.

The role of the state in mitigating favorable conditions for insurgency

At least four aspects of state capacity directly affect Fearon & Laitin's 'favorable conditions for insurgency' (Hegre, 2018). Fist of all, rebel groups can more easily organize and fight in "safe haven" territories where governments fail to physically access and project authority (P. Collier & Hoeffler, 2004; Fearon & Laitin, 2003). Herbst (2000) has shown how this phenomenon applies to several large and poor African countries with borders drawn by their former colonial powers, such as the Democratic Republic of the Congo (DRC). Here, the government tends to only control the core, while the hinterlands to a great extent are left ungoverned[5]. Second, the state's coercive capabilities by way of its military and police force are important tools that enable states to deter or repress potential dissidents and to ensure compliance with the law. When deciding whether or not to rebel, actors have to contemplate the risk of capture and survival on the one hand and governmental sanctions on the other. As such, smaller, poorly trained and bureaucratically less organized state forces incapable of identifying who the rebels are and their whereabouts should be less deterrent than the opposite.

The final two aspects relate to the state's level of bureaucratic and administrative sophistication, and its ability to extract sufficient revenue to accomplish their desired goals. Both elements are crucial for the aforementioned aspects, but in addition, administrative and extractive capacity matter particularly for governments' ability to develop and implement conflict-reducing policies (Mann, 1984; Tilly, 1985; Levi, 1988; Besley & Persson, 2009; Harbers, 2015). In order for the state to implement policies and exercise control, it needs both financial and human resources, and an effective, adequately budgeted state bureaucracy should be better equipped to curb resistance *and* to accommodate grievances through for example the provision of basic services. These services *could* be related to health care, education, and voting rights, but according to Huntington (1968), Fukuyama (2014) and the stateness-first argument, ensuring political order and security should be prioritized before introducing mass politics (Gjerløw et al., 2021).

[5]See e.g. Soifer (2012) for a discussion of this phenomenon in Latin America

Although state power is not homogeneous across its arenas (Soifer, 2008), Fjelde and de Soysa (2009, p.8) describe how the state's coercive, administrative and extractive capacities are mutually supporting and interlinked: "governments rely on revenue to invest in the military, police, and bureaucratic apparatus, which in turn allow them to accumulate power for further penetration and extension of state rule." Ergo, where high-capacity states are able to implement adequately financed conflict-reducing policies, states of lower capacity simply possess fewer policy options or lack the *incentives* to pursue proper strategies. Rossignoli (2016, p.428-29) further explains the relationship between the state's capacities and its role of mitigating favorable conditions for insurgency:

> *"State-capacity is the key factor in affecting both Opportunity and Motivation: the former through the political-military control by the government of the country; the latter by increasing the administrative structure (bureaucracy, tax collection, etc.) [...] On the one hand, when a territory is locally dominated by a minority group which is highly motivated to achieve self-regulation [...], state capacity operates in preventing conflict [...] mainly through its military, repressive dimension. On the other hand, when incompatibility arises over government, state capacity operates through all its dimensions, in affecting the political structure of the country. In this case, a stronger (i.e. more capable) state is expected to lower the probability of civil war onset."*

The state capacity model does therefore not negate the motivational approach, but instead, it emphasizes the state's invaluable role in maintaining political order by quelling both motivations and opportunities to rebel (K. Gleditsch & Ruggeri, 2010).

2.2 State Building and "Stateness-first"

Starting from the influential claim that "war made the state and the state made war" (Tilly, 1993), most theories on state building argue that no society can be peaceful, prosperous, or work satisfactorily without a political *order* (Olson, 1993; Bates, 2009; Fukuyama, 2011). From the earliest times to the modern era, they

hold, violence has been the norm, and it was not until it was constrained by states that large-scale development first saw the light of day (Pinker, 2011; North, Wallis, & Weingast, 2009; Kaplan, 2008). Because monarchs realized that their ability to finance and control war was key to survival, they invested in fiscal infrastructures and later in institutions, which provided the state with the ability to penetrate society more efficiently and implement their chosen policies (Mann, 1984; Soifer, 2008). Naturally, the imposition of political authority and institutions threatened the interests of recalcitrant actors, which required coercive subjugation of resistance. Thus, historically, statebuilding has often been a violent process (Newman, 2014).

While Weber (1918) famously regarded the state as a source of political order through its monopoly over the means of violence itself, Bates, Greif, and Singh (2002) depart from the Weberian perspective and views order as a characteristic of an equilibrium. From their perspective, private warfare needs to be banned and political order centralized before certain "specialists in violence" can ensure a balance of coercion by providing public security in exchange for revenues paid by those consuming the judicial services (see also Tilly, 1985; Olson, 1993). This frame of mind is also evident in Hobbes' famous Leviathan:

> *"For by Art is created that great Leviathan called a Commonwealth or State, which is but an artificial Man; though of greater stature and strength than the Natural, for whose protection and defence it was intended ... [without the State] there is continual fear, and the danger of violent death; and life of man solitary, poor, nasty, brutish, and short."* (Hobbes, 2008 [Introduction, XIII])

Fundamental to these theories is also the expansion of commerce. When societies, or 'natural states', moved away from their violent and chaotic historical trajectories towards a more stable order, violence did not disappear but was instead organized. When organized and employed (preferably as a threat rather than reality) to defend property rights, peoples' incentives to engage in productive activity increased, and consequently, "stationary bandits" as opposed to "roving bandits" allowed economic activity to bloom under their protection and collection of taxes (Bates, 2008; Bates et al., 2002; Olson, 1993; Tilly, 1985; North et al., 2009). The people remained in

control over the means of coercion, and the equilibrium became channeled towards the provision of law, order and protection of property and economic production. The *modern* state, as Mann (1984) sees it, adds bureaucracy and territoriality to the state's political power. His "infrastructural power" is thus the aspect of the state that determines how far its bureaucracy can reach to excert control.

Whether following the Hobbesian or Weberian solution to civil peace, there is probably some truth to be found in Walter Lippmann's observation that "There is no greater necessity for men who live in communities than that they be governed, self-governed if possible, well-governed if they are fortunate, but in any event, governed" (New York Herald Tribune, Dec. 10, 1963 as cited in Huntington, 1968, p.2). As such, at the very core of populations' security lies the origins and mechanisms of its autonomous powers, which are centralized and territorially bound (Mann, 1984).

This is not to say that decentralized systems *cannot* provide security, but rather that this type of security historically has imposed a trade-off between order and prosperity (Bates et al., 2002, p.611). Suchlike societies are often characterized by privately supplied coercion, limited specialization, lacking institutions, vertical loyalty networks, and a fragile peace (see e.g. Verweijen, 2018; Levitsky & Way, 2012). Many of the difficulties confronting such states stem from how they were created, and has both evolved from and resulted in disparate groups - making it challenging to evolve into "more mature entities" in which *other* nation-states rather than groups within the state were considered the true competitors (Kaplan, 2008, p.36, see also B. Anderson, 1991; Hobsbawm, 2012; A. Smith, 2010).

Although a simplified explanation, the main idea is that the affinitive power of identity, group allegiance and political order ultimately paved the way for modern governance, which could, at least in the European context, develop a normative political community, organizations, and institutions (Kaplan, 2008; Gjerløw et al., 2021). Summarized, a large number of 'sequencing-' or 'stateness-first' explanations contend that building effective institutions built on Weberian features of bureaucracy is a prerequisite for state capacity, and that suchlike constitutents have beneficial effects on a variety of outcomes, such as the introduction of democracy (Zakaria, 2007; Fukuyama, 2011, 2014; Huntington, 1968). How to "get to Denmark" is

therefore, according to this explanation, done by prioritizing the *state* before any *outcome*. Although not identical to state capacity, Zakaria (1997) points to the fact that historically, constitutional liberalism has led to democracy, while democracy has not seemed to bring about constitutional liberalism - but rather instability, clientelism and slow economic growth. As an example, he points to the world's de jure increasing number of democracies, while at the same time, numerous of these 'democracies' have no foothold in a framework of law and administration. Without such a foothold he conclude, "democracy itself has no peace-inducing qualitites" (Zakaria, 1997, p.37).

'Sequencing-' or 'stateness-first' explanations have come under criticism for among other things lacking empirical evidence. It is also hard to "tease the effect of state-first sequences apart from that attributable to correlated historical developments, such as industrialization, European imperialism, or institutional diffusion between neighboring countries" (Gjerløw et al., 2021, p.3). Despite this, and in part due to the limitations of the thesis, I will move forward with the widely held belief that political order and particular bureaucratic features need to be in place prior to introducing mass policies simply because controlling, extracting, and implementing conflict-reducing policies would be extremely hard without some fundamental bureaucratic components.

What is important to emphasize however, is that the very idea of a state capacity rating scale from strong to weak, or even failed, is contested[6]. The concept seems to have been generated solely on the basis of deviations from the ideal modern Weberian state, in which three distinct features are pivotal: monopoly of violence, territoriality, and legitimacy. Furthermore, there is a strong disposition among academics to conflate the Weberian modern state with the concept of a *state* in general. Consequently, Carment, Landry, and Samy (2014) and Wai (2012) emphasize that asking when, if, and how particular polities cross the imagined threshold of 'stateness' may exacerbate connotations to the Eurocentric unilinear evolutionist logic.

After London, Paris, Brussels and Lisbon partitioned land and drew colonial borders in Africa and South Asia, it quickly became evident that the European

[6]On failed states, see e.g. Herbst (2000); Rotberg (2003); Bates (2008); Wilkinson (2014)

notion of applicable geography did not fit the reality of the regions' demographics. The new borders were insensitive to complex matters, and the consequences are still felt in borderland communities of states like Sudan, Mali, DRC and on the Indian sub-continent, where colonialism forced ethnically framed disagreements to be resolved within artificial structures (Wai, 2012; Bates, 2009; Mishra, 2016; Michael, 2014; Newman, 2014). Furthermore, political order as described above has often been fully or partly absent in part because state leaders have not been provided with the incentives to invest in state capacity. Soifer (2008, p.240) has importantly noted that one way of resolving the ambiguities in the conceptual frameworks of state building is to clarify *what* aspects of the state are being built, which is what I attempt when restricting state capacity to three dimensions that are minimally necessary for the modern state to execute its core functions.

Despite conceptual disagreement and unfortunate connotations, state capacity still constitutes one of the most widely discussed and employed concepts in relation to armed conflict (Koren & Sarbahi, 2017, p.274). It is also likely to uphold its popularity due to core elements shared by the states at the weaker end of the spectrum, as governments' failing ability to penetrate society, exercise control and encourage economic growth poses severe development challenges and threatens livelihoods in the populations within their borders (see e.g. Herbst, 2000, p.424 and Mann, 1984, p.189). Due to limitations related to the scope of this thesis, I have decided to move forward with the conceptualization of the state as explained in this chapter, although acknowledging the problems induced by universalism.

2.3 Summary and Remaining Research Gaps

This chapter has outlined the prevailing trends in armed conflict and offered a review of three explanatory strands in the literature. While grievances- and greed-based interpretations are both "rebel-centric", I have explained how weak state capacity - manifesting in for instance poor public security provision, lacking incentives and/or capabilities to address grievances and conflict-reducing policies, as well as to enhance socioeconomic opportunities and to build strong, durable institutions - makes

states more vulnerable as rebels tend to exploit such environments. Thereafter, I emphasized the crucial role of the state when it comes to maintaining political order by quelling both motivations and opportunities to revolt. I have chosen to proceed with the state capacity model despite its unfortunate connotations because the majority of relevant research indeed conclude that weak states and intrastate conflicts, more often than not, go hand in hand.

However, as will be investigated thoroughly in the next chapter, the measures of state capacity used in quantitative analyses often rise important implications. When reviewing the state capacity literature, it becomes clear that theoretical and empirical findings oftentimes are inconsistent, and that there exist a multitude of theorized aspects *linked to* state capacity but that are dubious *measures* of the same concept. This is likely due to the fact that a multitude of capacity conceptualizations favor a highly disaggregated set of state functions, which in turn may drift from a core theoretical focus on the state's *abilities* to implement desired policies.

Accordingly, Hendrix (2010, p.273) concluded that "despite its importance for the study of [...] intrastate conflict, state capacity remains a concept in search of precise definition and measurement." I do by no means question previous research' importance when it comes to advancing our understanding of causes and dynamics of civil conflict, but I will nevertheless argue that as long as they continue to present different, and *conflicting*, notions and operationalizations of state capacity, the field is in need of a systematic validity-assessment and better approaches to modeling. Consequently, I will in the next section scrutinize state capacity's definitions, dimensions, and operationalizations, and their construct validity.

Chapter 3

Theoretical Framework: Unpacking State Capacity

Despite the increasing realization that state capacity is a fundamental ingredient for effective governance and domestic peace, the state capacity *concept* is utilized almost as an all-encompassing description of everything that has to do with the functioning and performance of the state. Under the most minimal Weberian definiton, the state is an organization aspiring to monopolize the legitimate use of violence within its borders. Over the last decades however, the state has repeatedly been brought back in, expanding our understanding of its role, function and ambitions. Encompassing institutions beyond the coercive, scholars of political science now find themselves "confronted with a spectrum of statehood" (Berwick & Christia, 2018, p.72), ranging from the minimum definition to the maximum, where the successful state is one where laws and institutions are just, meritocratic, democratic and liberal; where service delivery benefits all; where security and property rights are guaranteed; and where the economy, market forces and private sector can operate more of less freely.

Given this extensive array of understandings of statehood, it is no surprise that views on state building and state capacity diverge. Because different studies have been concerned with different outcomes, they have also focused on different sets of state functions - such as the ability to raise revenue, to ensure economic growth, to control territory, to protect human rights, or to provide democracy. Most of

suchlike aspects are indeed *aspects* of state capacity, but they are not necessarily *state capacity* as such. Furthermore, because the concept is both elusive and latent in nature, it can only be observed indirectly through operationalizations developed in a theoretical context. As a consequence, a myriad of competing and potentially complementary measures have emerged, and empirical efforts to disaggregate state capacity have induced ambiguous results. In his study, Hendrix (2010) found that the concept's underlying dimensionality is low, with more than 90% of the variance explained by three latent factors. Interestingly, Vaccaro (2020) moves further and emphasizes that strong correlations between measures should not be taken as proof of high interchangeability, and illustrates that the choice of state capacity measurement has a substantial effect on empirical results. These are worrisome findings considering the abundance of theoretical logics and available measurement options.

The following chapter deals with these linkages, and seeks to develop an understanding of which underpinnings of state capacity that actually are fundamental for the state to fulfill its most crucial functions and thus avoid civil conflict. In order to do that, I first review competing conceptualizations and measurements of state capacity. Second, I elaborate on measurement challenges and construct validity, focusing criticisms on two threats to inference. I also theoretically explain my own choices when it comes to selecting indicators. Lastly, I scrutinize in detail the three dimensions I argue are mutually interlinked and the indicators by which they are operationalized, emphasizing coercive, bureaucratic-administrative and extractive capacity.

3.1 Conceptualizing and Measuring Capacity

> *"The state has been brought back in. It now lumbers through the halls of academe, a great clumsy creature that no one quite knows what to do with."*
>
> (Geddes, 1996, p.1)

3.1.1 Conceptualization

Although it is widely accepted that state structures, abilities and behavior are critical to maintaining civil peace, efforts in dissecting and scrutinizing Geddes' (1996) 'clumsy creature' have a long way to go. In its various forms, higher state capacity has been linked to enhanced prospects for economic development and growth (e.g. Evans & Rauch, 1999), modified likelihood of being infected by conflict contagion (e.g. Braithwaite, 2010), decreased risk of experiencing civil wars and conflict recurrence (e.g. Fearon & Laitin, 2003; Sobek, 2010; Hegre & Nygård, 2014; Gates et al., 2016) and greater ability to implement and commit to peace agreements (e.g. DeRouen et al., 2010) to name a few.

Nevertheless, interpretation and usage of the state capacity concept varies considerably across the literature, and it is regularly being employed and measured interchangeably with closely-related terminology such as strength, fragility, failure, legitimacy, and institutions (Cingolani, 2013; Centeno, Kohli, Yashar, & Mistree, 2017). With such a broad array of concepts and measurement, it should come as no surprise that the empirical record is left confusing and at times conflicting. Without further theoretical and empirical development, 'capacity' has the potential to "cease to be a productive, analytical concept" and "becoming a classic residual variable called upon to explain unexpected outcomes given particular combinations of causal factors" (Centeno et al., 2017, p.4).

Despite the multitude of conceptual varieties, the majority of academics build their conceptualization of state capacity either implicitly or explicitly on Weber's depiction of a state as a political organization with the ability to produce and implement rules *vis-à-vis* a population within a restricted territory (e.g. Geddes, 1996; Besley & Persson, 2009; Soifer, 2012; Berwick & Christia, 2018). Thus, what state capacity most fundamentally refers to, is the ability to project and exercise *power* in order to "get things done" (Dahl, 1957). A common distinction has later been made to separate "power over" from "power to". However, as discussed by Lindvall and Teorell (2016), this distinction is hard to uphold in politics, as the state's power to implement policies and to achieve goals by large depends on the level of 'power over'

its people. Furthermore, in order to project any authority, the state needs to deploy *resources* toward *policy instruments* (Mann, 1984; Lindvall & Teorell, 2016). While *policy instruments* pertain to the state's efforts to achieve the compliance of mass publics, referred to by Bemelmans-Videc, Rist, and Vedung (1998) as "sticks" (coercion), "carrots" (economic incentives) and "sermons" (information/propaganda), *resources* represent all available assets that the state control and can utilize to improve the way it "coerces, bribes, and persuades" (Lindvall & Teorell, 2016).

What is conceptually agreed upon in the state capacity research agenda is moreoever that the state has a role in maintaining political order (Fukuyama, 2011; Huntington, 1968), and that state capacity pertains to the structures and abilities of state organizations (Skocpol, 1979). Or, more specifically: the state's abilities to ensure internal sovereignty, to deal with challengers either by repressive or accommodative means (see e.g. Muller & Weede, 1990; Hegre et al., 2001; Sobek, 2010; DeRouen & Sobek, 2016; Fjelde & de Soysa, 2009) and to being effective in implementing policies and extracting revenue within its territory (see e.g. Mann, 1984; Tilly, 2003; Thies, 2010; Gjerløw et al., 2021). In other words, scholars tend to think about state capacity in three ways: as the state's material capabilities, its effect on society, and its territoriality (Hanson & Sigman, 2020).

Another characteristic these conceptualizations share is their view of state capacity as an interactive process. Mann (1984) and Soifer and Vom Hau (2008) explore the 'infrastructural power' of the state and claim that the state's institutional capabilities to exercise control, penetrate civil society and implement policy choices within the territory are relational in nature and to a large degree shaped by relationships between state- and non-state actors on the one hand and between different state agencies themselves on the second (see also Berwick & Christia, 2018). This relational nature "allows analysts to move past debates that juxtapose state and society as opponents to examine the varied forms of their interaction" (Soifer & Vom Hau, 2008, p.222) and is important for the state to effectively initiate, coordinate and implement policies.

In order to arrive at a satisfactory conceptualization of state capacity applicable to this thesis, I have taken as starting point Mann's (1984, p.189) 'infrastructual

power' as the capacity of the state to penetrate society and to "implement logistically political decisions throughout the realm." Similarly, Skocpol (1985, p.16) argues that "sheer sovereign integrity and the stable administrative-military control of a given territory are preconditions for any state's ability to implement policies" and that "loyal and skilled officials and plentiful financial resources are basic to state effectiveness in attaining all sorts of goals." These definitions are beneficial in a number of ways, not least because they acknowledge the power aspect discussed above, and because they combine this power with territoriality, state reach, and monetary and administrative abilities to 'get things done' independently of *which* sets of decisions or outcomes they relate to (see also Migdahl, 1988; Geddes, 1996). Moreover, they are not affected by normative standards on how state power is exercised. I therefore agree with Hanson and Sigman (2020, p.3) in that it is better to conceptualize state capabilities in terms of the "material resources and organizational competencies internal to the state that exist independently of political decisions about how to deploy these capabilities".

Additionally, several scholars argue that it is more constructive to view the extractive, coercive, and administrative dimensions of state capacity as mutually interrelated and codependent rather than to disaggregate them, albeit they are conceptually distinct (Soifer, 2012; Lindvall & Teorell, 2016; Berwick & Christia, 2018; Hanson & Sigman, 2020). Accoding to Tilly (1985), it is the need for coercive capabilities that leads rulers to adopt tax systems and provide services. Fjelde and de Soysa (2009, p.8) state that "governments rely on revenue to invest in the military, police, and bureaucratic apparatus, which in turn allow them to accumulate power for further penetration and extension of state rule". Thus, raising revenue is not only fundamental to support all state activities, but it also encompasses capabilities that are imperative to a wider set of abilities such as the administrative task to reach populations, collect information, to monitor payments, and to possess coercive means that ensure compliance with citizens "who do not like to pay" (Harbers, 2015).

According to Levi (1988, p.32), compliance is voluntary in that citizens are obedient because they choose to - but at the same time, "it is quasi-voluntary because they will be punished if they do not and are caught". Furthermore, "a well-

organized, rule-following and autonomous bureaucracy staffed by meritocratically recruited, competent personnel" should be better at adapting and implementing any desired state policies (Gjerløw et al., 2021, p.11). Finally, the ability to get anything done fundamentally depend on the state's capacity to preserve its borders, to contain threats, and to maintain internal order.

I thus follow a conceptualization of state capacity that emphasizes the importance and codependence of the three dimensions that are minimally necessary for the modern state to execute its core functions - and thus avoid armed conflict by quelling both motivations and opportunities to rebel: 1) the state's coercive ability to deter and repel challenges to its authority with force, thus providing security and maintenance of internal order, as well as compliance with the law, 2) the bureaucratic and administrative sophistication necessary to collect, monitor and implement data, information and policy, as well as to reach its citizens and to conduct decision-making in an impartial manner, and 3) the extraction of revenue, in order to finance and support all other activities.

Yet, as Lindvall and Teorell (2016) has argued, this conceptualization also lends itself to assuming intentionality. A state's capacity to preform its core functions is only of explanatory value if the state actually desires to govern and/or to implement policies. I therefore restrict the analysis to the minimally necessary capabilities for contemporary states to achieve *intended* outcomes (though acknowledging that intention is hard to measure), and I argue that it is reasonable to assume that most states wish to avoid civil conflict[7]. Some might argue that this is somewhat contradictory, as 'avoiding civil war' is a policy choice and outcome, but considering the negative effect of conflict on state capacity (see e.g. Thies, 2015), I argue that avoiding intrastate conflict is absolutely pivotal for the state to *function* at all.

[7]Unless, of course, leaders profit from suchlike circumstances, but analyzing these particular cases is beyond the scope of this thesis.

3.1.2 Measurements

When it comes to measurements, the literature is generally divided into at least two camps, where state capacity is observed in terms of function rather than form. A common critique of 'state capacity' is the danger of tautology, or the 'blurring lines' between causes and outcomes, which is particularly evident in attempts to operationalize the concept (Cingolani, 2013). The first camp takes as point of departure the state's key role for societal outcomes, and contend that state capacity should be measured by what the state produces (outputs). Examples include income levels and economic development (e.g. GDP per capita: Fearon & Laitin, 2003); redistributive efforts and public goods and service provision like health, education, infrastructure and electricity (e.g. infant mortality, primary school enrollment, road networks and nighttime light emissions: Bustikova & Corduneanu-Huci, 2017; Thyne, 2006; Herbst, 2000; Koren & Sarbahi, 2017); and mechanisms for societal involvement in political decisions (e.g. measures of regime type like the Polity scale: Hegre et al., 2001; DeRouen & Sobek, 2004; Cederman et al., 2013).

The first camp has distinct measurement advantages when it comes to data availability. GDP per capita for instance, whether proxying economic development, grievances, state capacity or administrative complexity, is available for a large number of countries over an impressively long time frame (Hendrix, 2010). Measures of public goods outcomes such as literacy rates or the percentage of children enrolled in primary schools has the same benefit and can be collected from e.g. UNICEF, other UN institutions or the World Bank, but this type of data should at the same time be used cautiously, as the conditions under which they are measured are highly uncontrollable. They are furthermore generally hard to measure, they risk capturing other causal effects, and accurate numbers often depend on the state's ability to reach populations, collect information, and their incentives to report correct data.

Another serious disadvantage given priority to by for example Soifer and Vom Hau (2008) and Hanson and Sigman (2020) is that these approaches risk conflating the state's true *capacity* to implement such policies with the political decision to *prioritize* these functions. After all, the resources constitutive of state capacity

are different from the actual *use* of these resources. This makes it hard to capture differences in the ways in which power is exercised, or variation within states - which furthermore risks lumping together a fairly capable but minimal state (e.g. China) with a state of fairly low capacity (Soifer, 2012; Berwick & Christia, 2018). Moreover, measures of societal involvement in political decisions are deeply entwined with symptoms of political regimes, which is conceptually different than state capacity. Empirical evidence on the unequal territorial reach and performance of relatively democratic countries suggests that the capacity to get things done should not be equated with levels of democracy (Harbers, 2015; Giraudy, Luna, & Mistree, 2017).

In contrast, the second camp postulates that output-based measures are subject to normative standards on how state power is excercised and potentially influenced by exogenous events beyond the state's immediate control (e.g. Fukuyama, 2013; Soifer, 2012; Berwick & Christia, 2018; Hanson & Sigman, 2020). Instead, the attention is directed towards the more minimalist and normatively neutral definitions that are independent of political decisions on how to prioritize capacities. For Fukuyama (2013, p.354), state capacity consists of inputs, which he describes as resources and degree of professionalization in the bureaucracy and government officials. As Berwick and Christia (2018) emphasizes, physical resources are particularly amendable to quantification, and to this end researchers like Thies (2015, 2010), Lieberman (2002), Harbers (2015) and Besley and Persson (2009) rely on revenue data to illustrate that taxation is assosiated with property rights, the reach of the state, administrative sophistication, and of course, the ability to extract resources. Taxation data is commonly expressed as the ratio of state revenue to GDP, but another approach is to compare levels of income taxes to levels of taxes on international trade, as the former is expected to be more administratively complex to collext due to information gathering, state reach, and penetration (Soifer, 2009).

Furthermore, the second camp puts weight to the state's general abilities to *get things done* beyond its physical resources. State input therefore also includes autonomy, human capital, bureaucratic procedures, and ability to collect information. One argument is that state outputs depend on the state's original capacity to maintain autonomy, to coordinate policies and to gather information and thus

superior coercive capacity and bureaucratic-administrative sophistication, which is only possible to develop if it has generated resources by its extractive capacity. If the state is of low capacity, outcome preferences does not really matter:

> *"Where the state is weak, it cannot effectively carry out many kinds of policy, including those redistributing wealth, which are said to be at the heart of actors' preferences over regime outcomes."* (Soifer, 2013, p.3)

Autonomy, in the sense of a state's coercive ability to preserve its borders, protect against both external and internal threats, to maintain political order and to enforce compliance with the law is typically measured with data from the Wold Bank, Correlates of War or Stockholm International Peace Research Institute on military expenditures and the size of armed personnel and security forces. Coverage and reliability is generally good for most countries after the 1960s, but the relationship between military force and coercive capacity can be compared to the relationship between GDP per capita and administrative capacity in the way that there prevails widespread uncertainty about the true causal mechanism linking the two - which will be discussed in greater detail later.

On the other hand, general indicators of state functions capturing bureaucratic quality are available from a plethora of institutions, including World Bank, Political Risk Services Group, and the Quality of Government Institute. Moreover, different features of effective policy administration are commonly observed through measures of impartiality of state officials, meritocracy, popular observance of the law, tax collection advancement, and expert-coded measures of corruption levels to name a few. A potential drawback with expert-coded indicators is that these by necessity are subjective, and different perceptions of 'true' latent values, prior knowledge affecting assessments, or scale-ranging perceptions have the potential to put reliability at risk (Marquardt, 2020). When it comes to the state's ability to collect information, Brambor, Goenaga, Lindvall, and Teorell (2020) provide a measure of information capacity based on five information-gathering and -organizing activities. The World Bank tracks another aspect, statistical capacity, based on the state's ability to collect, analyze and disseminate data about its population and

economy.

Because state outputs and outcomes are products of the physical, human and informational resources that I argue are constitutive of state capacity, the conceptualization followed and measurements reviewed in this thesis follow the second camp. I wish to contribute to the enduring state capacity debate not by exploring *how* states perform specific functions or deploy resources to achieve certain political goals, but rather, independent of political institutions or regions, to scrutinize which capabilities that allow the state to take action.

Nevertheless, as rightfully noted by (Lindvall & Teorell, 2016), there are other ways for the state to promote growth and increase the stock of human capital that are unrelated to for example fiscal or bureaucratic capacity. One example is the provision of public education, which empirically suggests increasing state capacity in the long run. However, these measures are also likely to capture other causal mechanisms. How existing *input* measures of state capacity perform in terms of construct validity is investigated in the next section.

3.2 Theoretical and Empirical Challenges

When determining which and why countries are more, or less, prone to different types of violence, the point of departure is often the country-year unit of analysis in a large-N sample. From there, the researcher can investigate how national attributes operate in affect the propensity for armed conflict (Cunningham, D.E. and Lemke, D., 2014). Despite the superiority over its qualitative counterpart in terms of transparency, rigorousness and comparability, quantitative measures are by necessity standardized, which complicates linking latent concepts to observations. Consequently, not all empirical findings in otherwise valid and reliable inquiries reflect the theoretical logics they wish to have explored.

Thus, any attempts to measure state capacity must take into account the concept's *latent* nature. A latent concept is here taken to mean the result of a systematized formulation of a background concept conceived of in a theoretical context,

where no criterion is entirely adequate to empirically approach it in a direct way (Carmines & Zeller, 1979; Adcock & Collier, 2001; Gomide, Pereira, & Machado, 2018). One challenge of dealing with this type of concepts is therefore deciding upon the relations between the actually observable phenomena and their causes, and from there, seeking to measure the underlying concept based on theoretically derived hypotheses.

Two obvious, and perhaps the biggest, measurement issues when trying to measure state capacity in relation to civil conflict is data availability and endogeneity to the conflicts these measures are trying to predict. This is especially true for the states of lower state capacity, as these are often already mired in conflict, causing a reverse effect of decreasing state capacity, or they are lacking the minimally necessary capacities to execute the most fundamental functions of the modern state, which include reporting information and statistics. However, another challenge concerns construct validity and whether the indirect measures we *do* have in reality can be said to independently capture the theoretical construct of interest (Cronbach & Meehl, 1955; Blalock, 1979; Hendrix, 2010).

3.2.1 Construct Validity

As state capacity, validity too is a multifaceted concept. In its simplest form, validity brings in the idea of purpose and concerns the interpretation of observations arising from a specified procedure (Carmines & Zeller, 1979; Kirk & Miller, 1986; King, Keohane, & Verba, 1994). Validity traditionally touches upon especially three areas: the trustworthiness of internal causal relationships, whether and with what certainty findings are generalizable, and if the scores of an instrument measure the distinct concept. *Construct validity* deals with all of these questions simultaneously and shall in this thesis follow the definition suggested by Carmines and Zeller:

> *"the extent to which a particular measure relates to other measures consistent with theoretically derived hypotheses concerning the concepts (or constructs) that are being measured."* (1979, p.23)

Put differently, construct validity asks to what degree the operationalized variable

in an empirical model accurately reflects the variability among objecjects commonly arrayed on the underlying continuum to which the concept refers, and is particularly relevant "whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured" (Cronbach & Meehl, 1955, p.282). The general validity and quality of empirical research and claimed causal relationships should therefore be questioned, at best, if the measure itself 1) does a poor job capturing the theoretical construct, or 2) fails to discern between competing causal mechanisms (Hendrix, 2010). Although state capacity remains one of the most comprehensively discussed and employed concepts in political science, the preceding section disclosed an army of competing and sometimes contradicting definitions and operationalizations. Therefore, I will from here critically assess these operationalizations from both theoretical and empirical standpoints, based on the two aforementioned threats to inference.

3.2.2 Selecting Measures of State Capacity

Table 3.1 provides a list over common poor state performance (PSP) indexes. Indexing states according to their *performance* calculated from critical cut-off points became increasingly popular in the first decade of the 21st century, yet they are frequently criticized from numerous holds (e.g. Sanín, 2011; Grävingholt, Ziaja, & Kreibaum, 2015). One reason for this is that they tend to simplify the complex reality behind the decay of statehood to such an extent that they are of very limited use in crafting policies meant to counter state fragility or failure.

Table 3.1: List of common PSP indexes

Index (Source)	*Periodicity*	*Variables*
State Fragility Index (Center for Systemic Peace)	Yearly, 1995-2018	8
Fragile (previously Failed) States Index (Fund for Peace)	Yearly, 2005-2021	12
Bertelsmann Transformation Index (Bertelsmann Stiftung)	Biennial, 2003, 2006-2018	>50
Ibrahim Index of African Governance (The Mo Ibrahim Foundation)	Yearly, 2010-2019	>50
Index of State Weakness in the Developing World (Brookings Institution)	2008	20
Country Indicators for Foreign Policy Fragility Index (Carleton University)	Unknown, 2008, 2010-2012	>50

As point of departure for selecting which state capacity measures to scrutinize, I

purposely chose to exclude PSP indexes, in part because state fragility and failure indexes provide very little genuine information about performance orderings. Not only do they "distract attention away from analyses concerning the dynamics of state capacity" (Di John, 2010, p.10), but they also "only generate genuine information at the 'extremes' of their conceptual space, precisely where it is less useful and interesting" (Sanín, 2011, p.21). Furthermore, building PSP indexes necessarily requires having a conceptualization of what a state is and how it ought to function according to certain criteria, which generally has been done by prototyping. Fund for Peace's Fragile States Index for instance, assosicates governance indicators, regime type and state capacity indicators, and combine thresholds of economic growth, group grievances, security apparatus, state legitimacy, provision of public services and demographic pressure, amongst others[8] (Börzel & Risse, 2016). Unsurprisingly, some of these issues also relate to the state capacity concept when emphasizing the 'minimally necessary' abilities of contemporary states to carry out their 'core' functions, which generally are open to multiple interpretations.

Furthermore, the proliferation of performance labels - other examples include crisis, collapsed, shadow, and weak states - exclusively reflects the liberal view of war and violence. If there are multiple ways of 'falling apart' (Bates, 2008), a single conceptualization of fragility should be deemed inadequate, particularly when comparing the 'failure' of for example Yugoslavia and the Soviet Union with cases like DRC and Somalia. Indeed, when transforming social concepts into quantifiable values, we capture objects, not relations, and as a result, historicity not seldom becomes jeopardized. The final, and most important, reason to exclude PSP indexes from my analysis is that state capacity is conceptually distinct from state fragility or failure (Soifer, 2012). My aim is to scrutinize the construct validity of common *capacity* proxies and to examine its general underpinnings, not to identify the well-understood fact that Latin American or African cases more often than not fall short of the stateness commonly observed in East-Asia or Western-Europe.

Similarly, but unlike for instance Hendrix (2010), I have also chosen to avoid measures linked to the nature of state institutions for reasons described in the forego-

[8]See https://fragilestatesindex.org/indicators/

ing section. Although there exist an abundance of studies linking state *performance* to mechanisms for societal involvement in political decisions, I find it more constructive to focus on the capacities that exists within the state's organizational structures, and their territorial reach, independently of political decisions on how to deploy and allocate them. States like Russia, Singapore, China, and Qatar, of which all are considered fairly to very capable states, have for instance shown that the ability to avoid major civil violence does not depend on democracy. Although this link can be interpreted in relation to the prediction of Gurr (1970) and Hegre et al. (2001) that harshly authoritarian states and institutionally consistent democracies are both less likely to experience civil violence than semidemocracies, I argue that the effect is more likely captured by state *capacities* to ensure compliance, coercion, state reach, information gathering and control rather than *regime type* or mechanisms for societal involvement in politics itself. Indeed, governments that have not won over the "hearts and minds" of their people tend to compensate by investing in more coercive and repressive means.

The last two grouping of measures I have chosen to avoid relate to societal outcomes that are either beyond the state's immediate control, such as economic growth rates or income levels, or that depend on the state's level of prioritization. For the reasons already explained, I will from here on scrutinize indicators that relate to the three sets of capabilities that the state must possess to fulfill its most fundamental functions, nevertheless acknowledging that aspects of e.g. economic development and regime characteristics cannot be completely avoided. Here, what I seek to demonstrate following Hendrix (2010) and Hanson and Sigman (2020) is that state capacity is a multi-dimensional concept and thus unlikely to be properly captured in a single disaggregated measure.

Coercive Capacity

In order to grasp the dynamics of armed conflict within states, one has to, at least to some extent, scrutinize the driving factors of potential rebels. Considering the inherently militarized act of armed rebellion, the state's coercive capabilities has

occupied a privileged place in associated empirical studies (Hendrix, 2010). If adhering to the assumption that perceptions of relative power and opportunity cost relationships matter, the choice for rebels who consider whether or not to engage in costly conflicts should depend on the probabilities for winning in some form (Fearon, 1995; Cunningham, Gleditsch, & Salehyan, 2009; de Soysa et al., 2021). The model suggested by rational choice theory thus predicts that rebels factor the probabilities for capture, injures, imprisonment and death prior to the militarized act, which should at least in theory be affected by the state military's size, strength, and skills.

Because they by definition are fundamentally weaker than their government, rebels might choose to take up arms if they somehow perceive their government's military capacity to be weak. For instance, the shortcomings of the Nigerian army to exercise control and effective counterinsurgency (COIN) measures in the county's North-East has created a fertile environment for the emergence of not only jihadi groups like Boko Haram, but also local groups competing for influence. Furthermore, Harris (1970) points to the Cuban intelligence's assessment of the Bolivian army as poorly organized as the main reason for the guerilla's initial military superiority.

Coercive capacity has therefore regularly been defined according to the government's ability to physically deter potential challengers with force, to provide security and maintain internal order, and to ensure the quasi-voluntary act of compliance. States with limited coercive capacity typically lack sufficient means to provide security, limit rebel activities and to sanction, and they are less in control of borders, hinterlands, and safe havens (DeRouen & Sobek, 2004). Hence, this dimension is widely accepted as a fundamental dimension of state capacity and as a precondition for any further abilities (Weber, 1918; Skocpol, 1979; Mann, 1984; Tilly, 1985; Levi, 1988; Cingolani, 2013; Savoia & Sen, 2015; Lindvall & Teorell, 2016).

Typically, coercive capacity is operationalized by the number of armed personnel, security forces or by military expenditures with data from the Correlates of War Material Capabilities dataset (Singer, Bremer, & Stuckey, 1972), the Stockholm International Peace Research Institute (SIPRI), the World Development Indicators (WDI), or the International Institute for Strategic Studies (ISS). This operationalization is common to for instance Mason, Weingarten, and Fett (1999), DeRouen and

Sobek (2004) and Walter (2006), and studies generally find that more sophisticated militaries are associated with lower likelihood of conflict onset and higher likelihood of termination. Although few would doubt that "ceteris paribus, larger and more technologically advanced military forces are more effective" (Hendrix, 2010, p.277), the relationship between a state's coercive force and its coercive capacity is not straightforward (see also Soifer & Vom Hau, 2008).

From the perspective of construct validity, these measures are at least partly problematic. Larger armies should, at least in theory, deter potential rebels, and higher military spendings can increase military sophistication, surverillance capacity and repressive abilities, but at the same time, a large military force might be a symptom of increased threat perceptions (and thus compensation through investments in repression), insecurity, or war, and increased monetary priority attached to the military could just as well reflect corruption, patronage, or characteristics of regime type (Henderson & Singer, 2000; Gupta, de Mello, & Sharan, 2001; Conrad, Kim, & Souva, 2013). Furthermore, although states of higher coercive capacity may more easily maintain domestic order, some of the most peaceful countries in the world have either no or limited military[9].

Nevertheless, coverage and reliability for these measures are generally good after 1960. I include military spendings per 2018 constant US$ in population, military expenditure as a percentage of GDP obtained from SIPRI, and the number of armed forces personnel per thousand in the population from IISS and the Correlates of war (Singer et al., 1972). Expenditure as % of GDP should reflect the allocation of resources in an economy and how much of available output is devoted to defense, but it reveals nothing about the *level* of military expenditure. This is why a per capita measue in constant 2018 US$ is added (R. Smith, 2017). In addition, I include a measure of police personnel per 100,000 capita from the United Nations Office on Drugs and Crime. Of course, as critics have argued, these direct measures may capture the threat power of the government, but they reveal little about their

[9]According to the Global Peace Index' measure of peacefulness in 2020, Iceland and Costa Rica ranked 1 and 32 respectively. Both countries have limited military forces and no standing armies. Data retrieved from `https://www.visionofhumanity.org/maps/#/`

administrative ability to mobilize and to make coherent use of this potential.

Considering the potential issues with military spending and personnel, and because there might be other aspects of the state's coercive capacity that capture the theoretical construct, I also include some expert-coded measures. In order to capture the level of state presence and institutionalization in the territory, I firstly use Varieties of Democracy's (V-Dem) measure of the state's authority over its territory (v2svstterr), meant to "judge the extent of recognition of the preeminent authority of the state over its territory" (Coppedge et al., 2021, p.188). This variable is interpreted by for instance Ricciuti and Rossignoli (2017, p.2) to proxy state reach, "that simultaneously consider the multidimensional features of state capacity". Second, I include ratings from the Political Risk Services' International Country Risk Guide (ICRG) on law and order, measuring the impartiality of the legal system and the popular observance of the law on aspects like crime rates and effective sanctions (Howell, 2018). Law and order is measured on a 0-6 scale in which higher points reflect countries more in control of these aspects.

Third, I have included Hanson and Sigman's (2020) extension of the State Antiquity Index developed by Bockstette, Chanda, and Putterman (2002), who suggest that "an early territory-wide polity and experience with large-scale administration may make [...] for more effective government" [p.367]. Finally, I scrutinized the Bertelsmann Transformation Index' measure of the extent to which a state holds a monopoly on the use of force in its territory (Stiftung, 2020). This measure was excluded due to its non-representativeness, only covering "least developed countries" biannually since 2006. Table 3.2 summarizes the included indicators of coercive capacity.

TABLE 3.2: Indicators of coercive capacity

Indicator	*Source*	*Countries*	*Coverage*
Armed forces personnel/1000 capita	IISS, COW[a]	165	1991-2018
Military Expenditures as % of GDP	SIPRI	160	1991-2019
Military Expenditures/capita (constant 2018 US$)	SIPRI	159	1991-2019
Police Officers/100,000 capita	UNODC	97	2003-2017
State Authority over Territory	V-Dem v11.1	165	1991-2019
Law and Order	ICRG	134	1991-2019
State Antiquity Index	Bocksette et al. 2002	158	1991-2015

Notes: (a) Where the original source (ISS) reports country-year missingness but data on military personnel is available from the Correlates of War, I have subtituted ISS missingness for COW data (a total of 73 country-years)

Bureaucratic-administrative Capacity

If we again approach insurgents' considerations prior to resorting to rebellion, it is natural to think about the probabilities for capture. As Fearon and Laitin (2003, p.79-80) argue, "if government forces knew who the rebels were and how to find them, they would be fairly easily destroyed or captured. This is true even in states whose military and police capacities are low". This aspect has more to do with the bureaucratic and administrative capacity of the state than the coercive, as the ability to monitor, deter, suppress, control, and to gather information about the identities and whereabouts of their citizens "is a more direct determinant of the feasibility of rebellion than military capacity" (Hendrix, 2010, p.274, see also e.g. Fjelde & de Soysa, 2009). Thus, the state's ability to collect and manage information is of critical importance, perhaps *especially* where the terrain is favorable to rebels and characterized by rugged mountains and dense tropical forests (Herbst, 2000; P. Collier & Hoeffler, 2004; Soifer & Vom Hau, 2008; Buhaug, Gates, & Lujala, 2009; Hendrix, 2011). Summarized, "administratively weak states reduce the relative cost of organizing rebellion" (Fjelde & de Soysa, 2009, p.8), and such countries will have a harder time sanctioning, and preventing rebellion to occur in the first place.

This is consistent with what for example Skocpol (1985) and Fukuyama (2013) argue. For them, skilled and professional government officials are basic to state effectiveness, which is also recognizable from Weber's (1918) emphasis on the bureau-

cratically organized rational-legal structure of authority. Another argument for why bureaucratic-administrative capacity should matter for peace, points to the role of effective implementation of conflict-reducing policies, be it enhancing educational outcomes, health facilities, capital investments, or something else. If "co-opting political opposition and retain loyalty from key segments of society thorugh spending on political goods" is important for ensuring compliance (Fjelde & de Soysa, 2009, p.9, see also Levi, 1988; Taydas, Peksen, & James, 2010), the state must first and foremost possess the basic administrative infrastructure required to implement such policies (Hanson & Sigman, 2020). And, as noted by Gjerløw et al. (2021, p.11), a well-organized, rule-following bureaucracy "staffed by meritocratically recruited, competent personnel" should be better at adapting and implementing *any* policies.

Thus, the bureaucratic-administrative dimension of state capacity taps into a variety of aspects such as the professionalization of state agents, technical sophistication in terms of data collection, record-keeping and management, the monitoring and coordination mechanisms that are needed to effectively regulate economic activity and to reach populations across the state's territory, and lastly, the development of coercive institutions that can enforce the transfer of resources from the population to the state and impose sanctions if compliance is absent (Besley & Persson, 2009; Fjelde & de Soysa, 2009; Centeno et al., 2017; Harbers, 2015; Hendrix, 2010; Soifer, 2008; Hanson & Sigman, 2020). In turn, all of these aspects are important for the ability of the state to implement government's policy choices (Williams, 2020).

Although few would disagree with what the opportunity-model suggests - that low-quality institutions incapable of managing society reveal circumstances under which insurrection becomes more feasible - there is huge variation in the degree to which effective bureaucracy and administration is provided in-between Denmark and Somalia (Fukuyama, 2011; Börzel & Risse, 2016). Other features tap more closely to levels of corruption, court impartiality and government legitimacy and credibility, though such aspects at the same time are closely related to underlying principles of regime type and democracy, which assumes accountability, transparency, equal access to the political process and fair distribution and application of the law (Taydas et al., 2010). Nevertheless, one can expect a heavily corrupt and biased administra-

tion to be less effective, as people are more likely to be enabled governmental positions through patronage rather than ability, and because law enforcement could be compromised by opportunities of personal enrichment, thus decreasing the chances of effective implementation of state policies. It furthermore generates distrust and dissatisfaction among ordinary citizens, which could intensify grievances that are hard to either suppress or accommodate, especially in the absence of strong coercive and administrative abilities (Soifer & Vom Hau, 2008; Sobek, 2010).

In terms of measurement, there exist equally many operationalizations as there are aspects to a state's bureaucratic-administrative abilities. As previously discussed, a common way to measure bureaucratic and administrative capacity is by looking at outcomes of public goods and service delivery. However, a state could be in possession of the resources and administrative capacities necessary for providing such outcomes without actually prioritizing them. Second, measures like GDP/capita, infrastructural connections, health, and schooling may just as well capture other, closely linked concepts like economic development, grievances, and the nature of the political regime, which may be plausibly considered both a cause and effect of bureaucratic-administrative capacity. Thus, the endogeneity risks compromising analytical leverage, and when construct validity is considered, suchlike operationalizations are questionable at best (Hendrix, 2010; Hanson & Sigman, 2020).

Instead, ICRG's measure of bureaucratic quality should be more satisfactory from the perspective of construct validity, and is one of few measures that focus on the strength of the bureaucracy with relatively broad coverage (Howell, 2018). Assuming validity in measurement, this measure capture aspects more directly linked to the make-up of the state bureaucracy, and is measured on a 0-4 scale where high points are given to countries whose bureaucracies have the strength and expertise to govern without drastic changes in policy, that are somewhat autonomous from political pressure, and that have established mechanisms for recruitment and training. I also include ICRG's assessment of corruption in the political system, which is concerned with "actual or potential corruption in the form of excessive patronage, nepotism, [...] 'favor-for-favors', secret party funding, and suspiciously close ties between politics and business" (Howell, 2018). Corruption is measured on a 0-6 scale,

giving higher values to *less* corrupt bureaucracies.

From the V-Dem dataset, I have scrutinized two measures in particular. The first, Rigorous and impartial public administration (v2clrspct) assesses the extent to which public officials are rigorous and impartial in the performance of their duties. This variable focuses on "the extent to which public officials generally abide by the law and treat like cases alike, or conversely, the extent to which public administration is characterized by arbitrariness and biases" (Coppedge et al., 2021, p.175). Here, too, higher scores indicate more rigorous and impartial public administration. Although it could be argued that this measure is at least partly complementary to the included measures of corruption on the one hand and law and order on the other, I view this measure as capturing a different aspect, as it is less concerned with *legal* impartiality and favor-for-favours-corruption, and more concerned with norms and behaviours in the general public administration. The second measure assesses the level of administrative meritocracy (v2stcritrecadm), and asks to what degree appointment decisions in the state administration are based on personal and political connections as opposed to skills and merit. Higher points reflects higher levels of meritocracy. Despite the fact that at least the second aspect of construct validity - whether the measures are satisfactory discerning between competing causal mechanisms - might be at risk due to likely high correlations with measures of regime type, I contend that the constructs are equally important for state capacity in terms of policy implementation effectiveness.

Lastly, I have included three measures aimed at capturing the information-gathering capabilities of the states. Firstly, the World Bank's assessment of statistical capacity in which countries are scored against 25 criteria in the areas of methodology, data sources, and periodicity and timeliness with scores on a scale of 0-100. Using this measure, the World Bank seeks to estimate states' abilities to collect, analyze and disseminate high-quality data about its population and economy. Second, Brambor et al.'s (2020) information capacity is an aggregate index scored from 0-1 based on five indicators of information-gathering and -organizing activities, assessing different aspects of registers, statistical yearbooks, and state agencies' abilities to process information about the country's territory and popu-

lation. Third, from arguments made by for instance Soifer (2012) and Centeno et al. (2017), I include Hanson and Sigman's (2020) calculated measure of census frequency with data on country censuses provided by the U.S. Census Bureau, as this measure also reveals some level of territorial reach. All three measures are relatively direct, and should be considered adequate operationalizations of the construct they seek to measure.

Additionally, four other measures have come under scrutiny. Worldwide Governance Indicators' (WGI) measure of government effectiveness (D. Kaufmann, Kraay, & Mastruzzi, 2003) was excluded in part because of the criticism relating to their aggregation procedure and their sometimes questionable governance indices[10] which are unsatisfactory to both aspects of construct validity, and because WGI scores would be partly duplicative due to the set of overlapping indicators included in this thesis. Furthermoe, BTI's measure of basic administration was dropped because it assesses aspects to the administration that are both causes and outcomes of state capacity, thus putting the second construct validation feature at risk.

Finally, the World Bank's CPIA ratings of quality of budgetary and financial management, and quality of public administration were excluded because they are simply non-representative, as over 80% of the countries they provide information on with populations over 500,000 are characterized as either low- or lower middle income countries. I do at the same time recognize that the measures I *have* included may be prone to measurement errors based on subjective analytical perceptions, as some after all are expert-coded measures, but they are nevertheless more suitable in terms of construct validity and coverage representativeness than for instance the four mentioned above. Table 3.3 summarizes the included measures of this second dimension.

[10]For an overview over these indicies, see `https://info.worldbank.org/governance/wgi/Home/Documents`

TABLE 3.3: Indicators of bureaucratic-administrative capacity

Indicator	*Source*	*Countries*	*Coverage*
Bureaucratic Quality	ICRG	134	1991-2019
Corruption	ICRG	134	1991-2019
Rigorous and Impartial Public Administration	V-Dem	165	1991-2019
Admin. Appointment Decisions Based on Skills	V-Dem	165	1991-2019
Statistical Capacity	World Bank	129	2004-2019
Information Capacity	Brambor et al. 2020	67	1991-2012
Census Frequency	Hanson & Sigman 2020	164	1991-2015

Extractive Capacity

Because I for the purpose of this paper understand states as a set of societal-embedded governing institutions that are minimally tasked to deliver order and avoid civil conflict, it is relevant to have a look at the perhaps most fundamental aspect that allows the state to take action. Levi's (1988, p.1) motivating premise is that "[the] history of state revenue production is the history of the evolution of the state". The state, she argues, can only extend its rule if has the revenue to do so. Tilly (1985, p.172) shares this mindset, arguing that "variations in the difficulty of collecting taxes, in the expense of the particular kind of armed force adopted, in the amount of war making required to hold off competitors, and so on resulted in the principal variations in the forms of European states".

Some contributers to the civil conflict literature therefore adopt a fiscal sociology approach and define state capacity in terms of the ability to access the material resources that enable them to get things done, as it reflects both "how much autonomy the central state has from the actors from which revenue is extracted" (Mann, 1993, p.361) and the state's ability to "reach their populations, collect and maintain information, [and] possess trustworthy agents to manage the revenue" (Hanson & Sigman, 2020, p.8). Thus, the ability to extract should inform all other dimensions of state capacity (see e.g. Bates, 2009; Besley & Persson, 2009; Cingolani, 2013; Herbst, 2000; Berwick & Christia, 2018). As Lieberman (2002, p.92) points out:

> *"Tax collection is ultimately the product of policy making, the monitoring of economic activity, the administration of complex laws, and judicial and punitive enforcement. For scholars, varied levels of tax revenues reflect variations in these state processes."*

The most efficient methods of extracting according to Soifer (2009), such as the national income tax, requires costly penetration of society through appratus like for example censuses, registers and other elements referred to in the foregoing section. States lacking this administrative capacity may turn to less lucrative but easier-to-collect trade taxes, whose enforcement is limited to border regions (Soifer, 2009; Besley & Persson, 2009; Berwick & Christia, 2018; Hanson & Sigman, 2020). This thesis does not follow the assumption that a state that is capable of extracting tax revenue necessarily also can exercise coercive power or that this is the definitive determinant of state capacity in general, however. State power is not homogeneous across its arenas, but nevertheless, a "well-organized" government with adequate budgets should be more equipped to pursue its goals (Skocpol, 1979; Herbst, 2000; Bates, 2009; Fjelde & de Soysa, 2009; Thies, 2010).

However, the literature is divided on how extractive capacity relates to peace. While for instance Buhaug (2010) finds "compelling empirical evidence" that states of greater ability to extract resources are significantly better able to maintain peace, Thies (2010) finds that revenue generation is not significantly related to civil war onset. Furthermore, extraction is highly correlated with and likely endogenous to economic development (Berwick & Christia, 2018). Still, direct measures of extractive capacity are attractive and generally have good coverage after 1970.

One widely used revenue-based indicator of the size and capacity of the state is *total taxes/GDP*, which is available from a number of sources such as the World Bank, IMF, and the International Centre for Tax and Development (ICTD). Two important shortcomings of this indicator relate to the fact that levels of tax revenue collection can reflect a *policy choice* rather than actual *capacity*, and that it fails to distinguish between the *types* of taxes that are extracted, ignoring differences in administrative sophistication. As noted by Hendrix (2010), countries as diverse

as Algeria, Lesotho and Sweden all had 1980-2002 average total taxes/GDP values near 31%, which suggests that they are comparable in terms of extractive capacity. I nevertheless include tax revenues as a proportion of GDP to capture overall extractive capacity, but in addition, I control for the second pitfall and include data from ICTD and the World Bank on the proportion of tax revenues - rather than as a proportion of GDP - income taxes/total taxes and taxes on international trade/total taxes. Here, I expect that a greater proportion of the former captures higher levels of both extractive and administrative capacity, since the latter is administratively easier to collect.

Following e.g. Thies (2010), Buhaug (2010) and Fjelde and de Soysa (2009), I also include a measure of "relative political capacity" (RPC), which assesses the amount of tax collected relative to an estimated expected amount of revenue (Organski & Kugler, 1980; Fisunoglu, Kang, Arbetman-Rabinowitz, & Kugler, 2011; Kugler, 2018). This measure compares a state's efficiency at extracting resources compared with other states with similar endowments, and addresses especially two limitations with the previous measures: it controls for wealth effects, and it accounts for factors that affect tax collection by modeling "the proportion of production flowing through convenient tax handles (mining, exports) and the demand for public expenditures" (Hendrix, 2010, p.279). In other words, RPC does not only capture taxes extracted directly from their populations, but it also account for the presence of revenue sources that do not depend on controlling people - so that "the actual extracted taxes capture the level of penetration of the society by the state" (Fjelde & de Soysa, 2009, p.12). States that collect tax revenues close to the predicted values will have RPC values near 1, while a state that collects double the predicted amount will have an RPC value of 2.

The measure's most significant construct validation-drawback is that it is potentially modeling the shape of the economy. Furthermore, countries that derive significant revenue from non-tax sources could have less incentives to access societal resources, and will not necessarily have strong incentives to invest in other dimensions of state capacity. This problem may cause the measure to be biased downwards, due to its close connection to theories related to the resource curse -

suggesting that countries that are rich in lootable resources can undermine state capacity because these resources are harder to control and tax, rebel opportunities to organize should be greater, capital is redeployed to the war, and natural resource wealth can reduce the need to cultivate a broader tax base (and the institutions associated with it) and increase corruption and economic mismanagement (see e.g. de Soysa, 2000; P. Collier & Hoeffler, 2004; R. Snyder & Bhavnani, 2005; Thies, 2010; Taydas et al., 2010; Wegenast & Basedau, 2014; Ross, 2015).

For these reasons, I also include a measure of natural resource dependence: natural resource revenue as a proportion of GDP, expecting that *higher* levels of resource revenue correspond to *lower* levels of both extractive and bureaucratic-administrative capacity (Newman & DeRouen, 2014). This indicator is calculated using ICTD's "total non-tax revenue" as the difference between total revenue and total non-resource tax revenue. The variable thus includes *all* types of non-tax revenue, including resource revenue, but acts as a useful proxy for resource wealth because natural resource revenue explains the vast majority of the variation in "total non-tax revenue" across countries (Prichard, Cobham, & Goodall, 2014). Lastly, I add V-Dem's expert-coded assessment of fiscal capacity (v2stfisccap), capturing the extent to which the government is able to fund itself from taxes of greater complexity (Coppedge et al., 2021). I also investigated World Bank's CPIA rating of revenue mobilization efficiency, but for the same reasons as with other CPIA measures, this was excluded. Table 3.4 lists the included measures of extractive capacity.

Table 3.4: Indicators of extractive capacity

Indicator	*Source*	*Countries*	*Years*
Total Tax Revenue as % of GDP	ICTD, IMF, World Bank	164	1991-2019
Taxes on Income as % of Taxes	ICTD, World Bank	134	1991-2019
Taxes on International Trade as % of Taxes	ICTD, World Bank	134	1991-2019
Relative Political Capacity	Kugler & Arbetman 1998	164	1991-2018
Natural Resource Revenue as % of GDP	ICTD	152	1991-2018
Fiscal Capacity	V-Dem	165	1991-2019

3.3 Summary

The most central question in this thesis is whether existing, disaggregated operationalizations of state capacity accurately and independently capture the theoretical concept, or if an aggregate measure of clustered characteristics that relate to core state functions is empirically preferable, particularly in relation to intrastate conflict causes and dynamics. In an attempt to answer this question, I have in this chapter outlined common conceptualizations and measurements of state capacity, and argued that especially three dimensions are interrelated, codependent, and hard to separate in practice, although they are conceptually distinct. The chapter furthermore includes a critical assessment of the indicators by which the dimensions are operationalized on the basis of construct validity, and a discussion of the bigger measurement issues when attempting to measure a latent variable like state capacity.

In summary, I have argued that in order for a state to secure internal peace and to implement any kinds of policies, it needs to be in possession of 1) revenue, 2) the coercive means required to ensure compliance and to contain threats throughout its territory, and 3) the administrative sophistication enabling it to reach and monitor populations, collect information, conduct decision-making in an impartial manner, and to implement desired policies. I refer to these dimensions as extractive, coercive, and bureaucratic-administrative capacity respectively. If a government lacks any of these capabilities, it will have a harder time curbing both opportunities and motivations to rebel, thus making it more vulnerable to civil conflict.

Focusing upon the key dimensions that enable the state to act provides the basis for a measurement strategy that steer clear of conflating state capacity with other closely-related terminology and normative outcomes, and that avoid disaggregating the concept into ever-smaller functional roles contributiing to ambiguous results. Thus far, I have theoretically explained the strengths of an aggregate state capacity approach solely focusing on the abilities that the state controls independently of how it chooses to deploy its resources. The next chapter outlines the methods used to investigate state capacity empirically, and accounts for the methodological implications that have appeared along the way.

Chapter 4

Data and Methodology

The preceding discussion theoretically demonstrated how highly disaggregated measures regularly shift the focus from states' ability to implement desired policies and instead conflate state capacity with either other, conceptually distinct state features, the political decision to prioritize them, or certain state functions that are only relevant to the researcher's particular inquiry. It furthermore identified 20 distinct indicators and assessed each measure according to two aspects of construct validity: whether the measures theoretically correspond to the concept, and whether they successfully exclude competing causal mechanisms. Each has merits and drawbacks.

This chapter develops a quantitative research design aimed at bridging the gap between the theoretical framework and the empirical data, and elaborates on the methods used to statistically test whether coercive, bureaucratic-administrative and extractive capacity in fact can be empirically distinguished. The first section elucidates the selection criteria and spatiotemporal scope of the collected data. The second scrutinizes the data structures more thoroughly, and accounts for the procedures and decisions made along the way. The third explains the method used to test the aforementioned assumption and to construct a new measure of the presumed concept of state capacity, namely principal factor analysis. Finally, I discuss methodological challenges, with a main focus on missing data and the multiple imputation approach.

4.1 The Quantitative Approach

For decades, there has prevailed an extensive back-and-forth banter among social science researchers regarding the differences between and advantages of the two approaches commonly denoted "qualitative" and "quantitative". Quantitative researcher Fred Kerlinger has for instance famously claimed that "There's no such thing as qualitative data. Everything is either 1 or 0" (Miles & Huberman, 1994, p.40). As a response, Donald Campbell replied that "all research ultimately has a qualitative grounding" (as cited in ibid.). Although the two traditions differ in terms of style and technique, their overarching goal is to design research that test certain assumptions, and eventually generate valid inferences on the basis of empirical information about social life (King et al., 1994; Mahoney & Goertz, 2006). In order for *any* research to be scientific, researchers have to present its procedures explicitly and transparently, evaluate its methods, and consider its conclusions' uncertainties in such a way that externals can scrutinize its methodology and thus evaluate its quality (see e.g. King et al., 1994; D. Collier & Mahoney, 1996; Moravcsik, 2014).

The qualitative approach is particularly useful when adopting a causes-of-effect approach, in circumstances where fine-grained case knowledge is sought (e.g. for policy informing purposes), as extra leverage in quantitative analyses, or in comparative studies seeking to match, contrast, and/or strengthen causal inferences in small-N designs either between cases or within a single case (Bennett, 2004; Mahoney & Goertz, 2006; Checkel, 2008). In large-N quantitative studies on the other hand, the objective is normally to uncover causal relationships that are statistically generalizable beyond the sample and/or to predict future events. While the qualitative approach can offer important assessments of context-specific case peculiarities, the quantitative alternative is generally accepted to outperform the former in terms of comparability, transparency, rigorousness and fitness for purpose in larger analyses. Statistical techniques are however less able to weight impacts of covariates differently and are thus more standardized, which has concerned researchers regarding a "lack of historicity" (e.g. Skocpol, 1979; Sanín, 2011).

Another concern, which alone serves as a justification for this quantitative in-

vestigation and which was dealt with thoroughly in the previous chapter, concerns construct validity. Research on the state capacity-intrastate armed conflict nexus can only be relevant, reliable and systematically unbiased if data and approaches to modeling are enhanced. Moreover, righteous policies can only be developed if trend-estimates across time and space are coherent and, as far as possible, representative. I have chosen the quantitative approach in an attempt to contribute to the field with a potentially more valid and reliable measurement of state capacity, although being aware of the difficulties and imperfectness of statistically measuring a social construct, in which "measurement error is probably the rule rather than the exception" (Clayton, 2014, p.36).

4.1.1 Selection Criteria

My main objective is to construct and test whether an aggregate measure of state capacity is empirically (and theoretically) preferable over the prevailing disaggregated measures, particularly when it comes to predicting intrastate armed conflicts in cross-national comparative research. As demonstrated in the foregoing chapters, the amount of quantitative studies linking state capacity to intrastate armed conflict dynamics is vast, and the operationalized indicators proxying state capacity equally so. I have therefore, based on theorized construct validity, sought to select the "best" subset of state capacity predictors. In doing that, certain selection critera have been fundamental.

Spatiotemporal scope

The analysis is based on the country-year unit, and the main observation of interest is a state's ability to implement desired policies and carry out its key functions in a particular year. A state is defined by N.-P. Gleditsch et al. (2002, p.619) as "an internationally recognized sovereign government controlling a specified territory, or a non-recognized government whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory." The selection of states to be included in the analysis is largely based on Gleditsch &

Ward's list of independent states (1999), but with populations $\geq$ 500,000 in 2018. A global sample is desirable when the aim is to generalize, and is relevant in this analysis because the core functions of modern states are the very same regardless of geographical regions. However, it might be questionable to compare levels of state capacity in Sub-Saharan Africa where intrastate conflicts occur regularly, with Northern Europe, where civil conflicts rarely to never see the light of day. Unfortunately, the relationship between state capacity and armed conflict is plagued by endogeneity, and this implication is discussed in more detail in section 4.4.

Microstates and regions whose government is internationally disputed are thus excluded by default. This is done because of practical matters, as I want the retained data to be as consistent across data sources as possible. Malta is nevertheless included because its population was close to the 2018 threshold and data availability adequate. Unfortunately, indicators and complete data on developing and/or colonial states prior to the 1980s are not readily available from a large number of sources - which has limited the data collection to at least partially be driven by data availability, as is often the case for cross-country quantitative studies of this kind (Hendrix, 2010; Nassiri, Lovik, Molenberghs, & Verbeke, 2018). North Korea was excluded due to a high proportion of missingness across data sources.

Although some sources, like the Correlates of War, Stockholm International Peace Research Institute, V-Dem and the World Bank, provide relevant and comprehensive data on most of these countries for decades prior to 1991, other sources such as ICRG have a more restricted temporal and spatial domain. I have chosen to restrict the analysis' temporal coverage to 1991-2019 of three reasons in particular. Firstly, the majority of today's sovereign states received their independence either by or in 1991. Second, data availability, global coverage, and quality is generally remarkably better from this decade. Third, I keep clear of countries that ceased to exist as a result of the Cold War and thus have no effect on contemporary state capacity relationships. A comprehensive list of the included countries as well as their geopolitical region (as categorized by Teorell et al., 2021), temporal coverage and the min/max number of available indicators per country can be found in Appendix Table A.1. It would be optimal to start the time series in the 1960s, when intrastate

armed conflicts had become the most frequent conflict type (Figure 2.1). However, biased missingness and data accuracy would likely have been issues with even more pressing implications. Starting the analysis in 1991 should still be satisfactory to cross-national studies of state capacity and intrastate conflicts set in the post-Cold War and -colonial era.

Conceptual fitting

The indicators that ended up being included in the analysis are, to the best of my knowledge, the most conceptually fitting, frequently employed, and recognized measures relating to the three identified dimensions of state capacity. Some of them are of course more closely related to one dimension than the other, but they are nevertheless closely interrelated and, at least in theory, interdependent. In a perfect world, there would exist indicators that would be fully capable of capturing each theorized dimension independently, but unfortunately, that is not the case. Thus, I can only compensate and make the best out of the data that exists by carefully evaluating the merits and pitfalls of each indicator, and subsequently collect the ones that correspond *the closest* to the conceptualized construct and that are *the least* connected to competing causal mechanisms.

Furthermore, in order to achieve as much representativeness and consistency across indicators as possible, the included measures are obtained from datasets that provide more or less global coverage. As such, the World Bank's CPIA indicators were excluded. There are especially four types of indicators I have sought to avoid. The first type consists of indicators that are closely linked to economic development, like economic growth, income, and redistribution efforts. The second captures societal/development outcomes such as infrastructure, health and education - which in themselves depend on economic development and political prioritization. The third covers political features and regime characteristics like democratic accountability, political trust and election of policymakers. These types of measures might indeed correlate highly with state capacity and be important for the general 'performance' of the state, but as argued in the last chapter, they reflect conceptually

different aspects of the state and also depend on political prioritization rather than the material inputs and organizational competencies internal to the state. Economic development is also largely affected by shocks exterior to the state's control.

The most "problematic" included indicators in terms of this (despite the ones capturing military capacities, as already accounted for) might be corruption and natural resource wealth, which arguably not directly reflect administrative or fiscal capabilities (Hanson & Sigman, 2020). Nevertheless, a number of studies provide evidence that the two undermine state capacity both separately and in combination, e.g. because personal enrichment and an abundance of primary commodity hamper incentives to invest in other dimensions of state capacity (P. Collier & Hoeffler, 2004; Bates, 2009; Taydas et al., 2010; Ross, 2015). Hence, I include both.

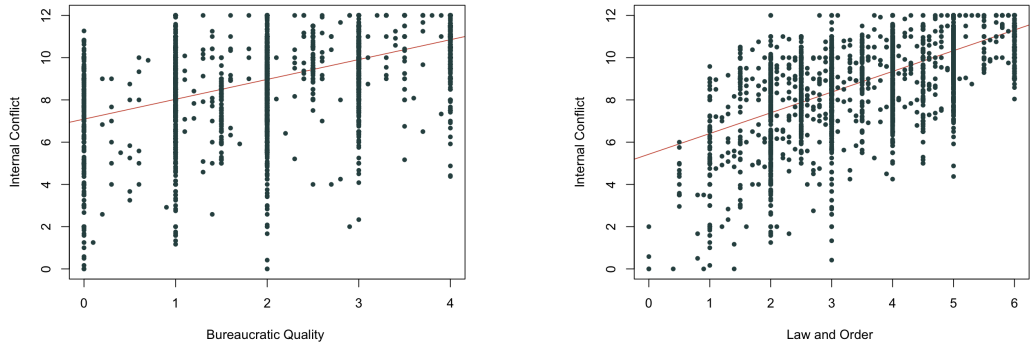
The fourth group of indicators I have steered away from accounts for PSP-indexes on the one hand and broad aggregate measures on the other. First of all, a great number of these measures are constructed from indicators that relate either to economic development, societal outcomes, or regime characteristics (e.g. Fragile States Index and WGI's Governemnt Effectiveness). Second, PSP indexes provide limited to no genuine information about performance orderings. Third, some include the same component indicators that are included in this thesis, leading to duplication. I do however include a measure of information capacity (Brambor et al., 2020) and a measure of statistical capacity from the World Bank. Both are expert-assessed aggregate measures, but are included because they are concerned with specific bureaucratic functions and distinct abilities not covered by other measures.

A mixture of subjective and objective measures

In order to overcome the fact that both types of measurements have flaws, I have sought to include a balanced mix of *subjective* and *objective* indicators. Subjective measures of state capacity, i.e. expert-coded assessments, are particularly prone to two sets of troubles. First of all, they are more prone to endogeneity problems because they are highly sensitive to information regarding history, economic development, unrest and conflict. As noted by Hendrix (2010, p.278), the endogeneity

of explanatory variables to conflict is not unique to expert assessments, but indeed because they are subjectively measured, the problem is likely to be magnified. For example, although ICRG has an isolated variable that measures internal conflict, this variable is highly correlated with their expert-coded measures of bureaucratic quality *(r = .51)* and law and order *(r = .64)*. Figure 4.1 presents this relationship graphically. Furthermore, unrest and violent clashes that fall short of the conflict threshold might also affect the coding on bureaucratic quality and law and order.

FIGURE 4.1: Scatterplot of internal conflict, bureaucratic quality and law and order



Another pitfall of subjective measures is that these measures are likely to be affected by different perceptions of scaling thresholds and judgements of 'true' latent values, which has the potential to put reliability at risk (Marquardt, 2020; Vaccaro, 2020). Although it is plausible to assume that rating divergences are lowest at the clear-cut extreme ends, intermediate levels of any expert-coded measure should be expected to be at least partially systematically biased - *especially* when coding is based on contemporary perceptions of distant points in time (Hanson & Sigman, 2020).

On the other hand, objective measures are those based on raw data, such as tax revenue, armed forces personnel and military spending. These measures can also be problematic, first and foremost because they might represent political decisions on how to collect and deploy resources. Also linked to the inherent challenge of endogeneity, states with larger armies or military spending could just as well be states mired in conflict as opposed to strong, coercive states. In that case, these measures would represent political preferences and prioritization rather than capacity per se.

4.2 Data Description

Altogether, I employ 20 different indicators of state capacity related to three key dimensions (Table 4.1). The indicators are drawn from multiple sources and span 29 years (1991-2019) for up to 165 countries in a given year, with a maximum of 4,722 country-year observations and 72,919 data points in total. In 96% of country-years, at least 10 indicators are available, and the median number of indicators per country-year is 16. A qualitative description of all indicators can be found in Appendix Figure A.1.

Table 4.1: Descriptive statistics

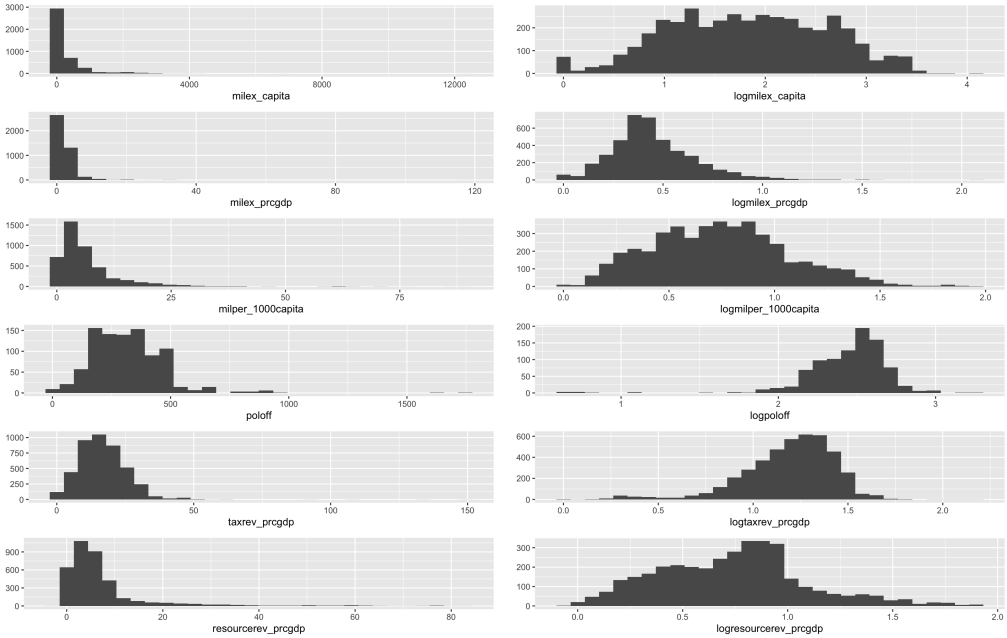
Variables	*Mean*	*Std.Dev.*	*Min*	*Max*	*Obs*	*Miss. obs*	*Ctry.*	*Years*
Armed Forces Personnel per 1000 capita	6.51	7.40	0	89.389	4427	295	165	1991-2018
Military Expenditures as % of GDP	2.25	2.88	0	117.35	4135	587	160	1991-2019
Military Expenditures per capita	263.48	522.65	0	12330.5	4135	587	159	1991-2019
Police Officers per 100,000 capita	318.65	156.79	3.15	1750.19	923	3,799	97	2003-2017
State Authority over Territory	91.24	10.94	33.6	100	4722	0	165	1991-2019
Law and Order	3.72	1.37	0	6	3791	931	134	1991-2019
State Antiquity Index	.69	.16	.24	1	3875	847	158	1991-2015
Bureaucratic Quality	2.16	1.12	0	4	3791	931	134	1991-2019
Corruption	2.82	1.26	0	6	3791	931	134	1991-2019
Rigorous and Impartial Public Administration	.37	1.49	-3.75	4	4722	0	165	1991-2019
Admin. Appointment Decisions Based on Skills	.43	1.08	-2.44	3.26	4611	111	165	1991-2019
Statistical Capacity	67.40	16.35	16.6667	98.8889	2007	2,715	129	2004-2019
Information Capacity	.65	.15	.24	1	1455	3,267	67	1991-2012
Census Frequency	.97	.38	.140845	2.32558	4045	677	164	1991-2015
Tax Revenue as % of GDP	16.67	8.82	.000125	149.284	4301	421	164	1991-2019
Taxes on Income as % of Taxes	36.01	17.12	0	130.54	2761	1,961	134	1991-2019
Taxes on International Trade as % of Taxes	11.58	14.04	-18.4962	79.1209	2762	1,960	134	1991-2019
Relative Political Capacity	.99	.39	.011937	3.37442	4474	248	164	1991-2018
Natural Resource Revenue as % of GDP	6.84	8.89	-2.3517	83.5436	3580	1,142	152	1991-2018
Fiscal Capacity	.91	1.22	-3.04	2.93	4611	111	165	1991-2019

4.2.1 Log Transformations

As can be identified by Table 4.1, six of the indicators are heavily skewed. For instance, military expenditures per capita has a mean value of 263.48 (e.g. Chile 2009 and Romania 2019) and a maximum value of 12330.5 (Kuwait 1991). In order to minimize skewness and map distributions closer to normal, I have logged these

variables. Figure 4.2 portrays the overdispersion by comparing the variables' original right-skewed distribution with the log-transformed versions.

FIGURE 4.2: Overdispersion



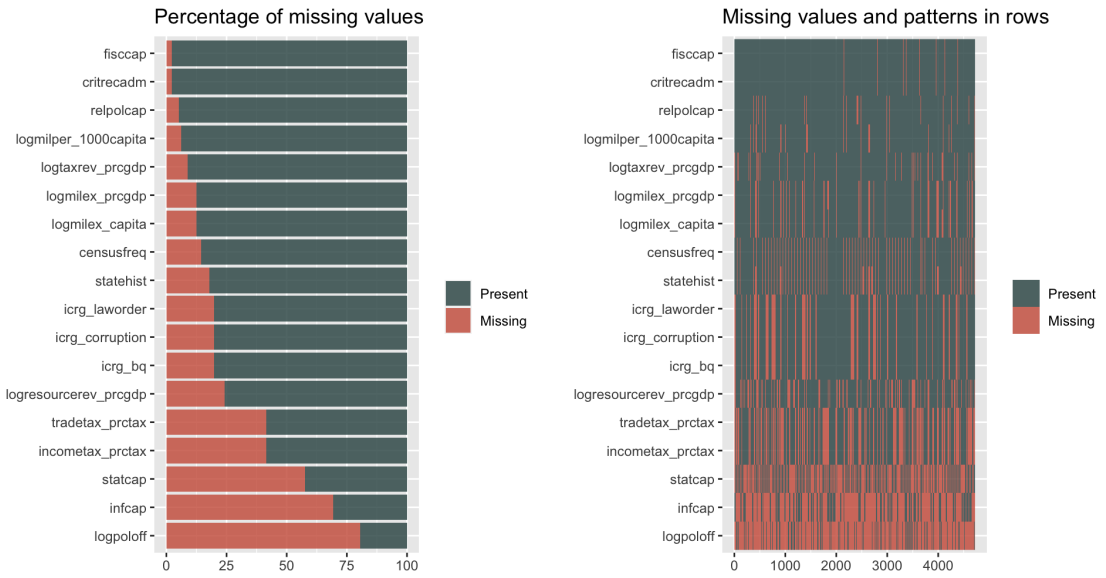
4.2.2 Missingness

Furthermore, all variables except for two contain missing values. Figure 4.3 exhibits the variables' missingness magnitude and patterns, and illustrates that the variables with the highest proportions of missingness are police officers per 100,000 capita, information capacity, statistical capacity, income tax as % of total tax, and taxes on international trade as % of total tax. In multivariate "time-series cross-section" data, the missingness phenomenon is not uncommon, and can be dealt with in various ways (Honaker & King, 2010).

One approach to analyzing data with missing values is referred to as complete case analysis, and includes listwise- and pairwise deletion (Nassiri et al., 2018). These methods deal with the phenomenon by removing all subjects with missing values so that the analysis only utilizes the actually observed data, or completely observed pairs in e.g. a covariance matrix. However, in my case, either of these techniques would lead to an extreme loss of information. Furthermore, because my

dataset only contains 115 sets of complete country-year observations as opposed to the possible 4722, and because the observed pattern of missing data is not missing completely at random, both approaches would severely bias estimates (Rubin, 1976).

FIGURE 4.3: Missingness magnitude and pattern



Another approach that has become increasingly common is to consider imputation, where missing values are replaced by imputed values (Honaker & King, 2010; Zhang, 2016). Among numerous methods developed for imputation, two of the most common include single mean/mode imputation and multiple imputation in the sense of Rubin (Rubin, 1976). Single mean/mode imputation is a fast and simple method for dealing with missing values that avoids a significant loss of information, but because values are imputed at the center of the variable's distribution, it necessarily results in an artificial reduction in variability. Additionally, the method shrinks standard deviations and compromises correlations between variables (Zhang, 2016).

On the other hand, multiple imputation (MI) uses the distribution of the observed data to estimate multiple values reflecting the uncertainty surrounding the true value (Johnson & Young, 2011). The idea behind MI is to extract as much relevant information from the observed data as possible via a statistical model, to impute multiple values for each missing cell, and subsequently to construct a certain number of "complete" datasets. In each dataset, the observed values remain

the same, while the imputed values vary depending on the estimated uncertainty in predicting each missing value (Honaker & King, 2010, p.561).

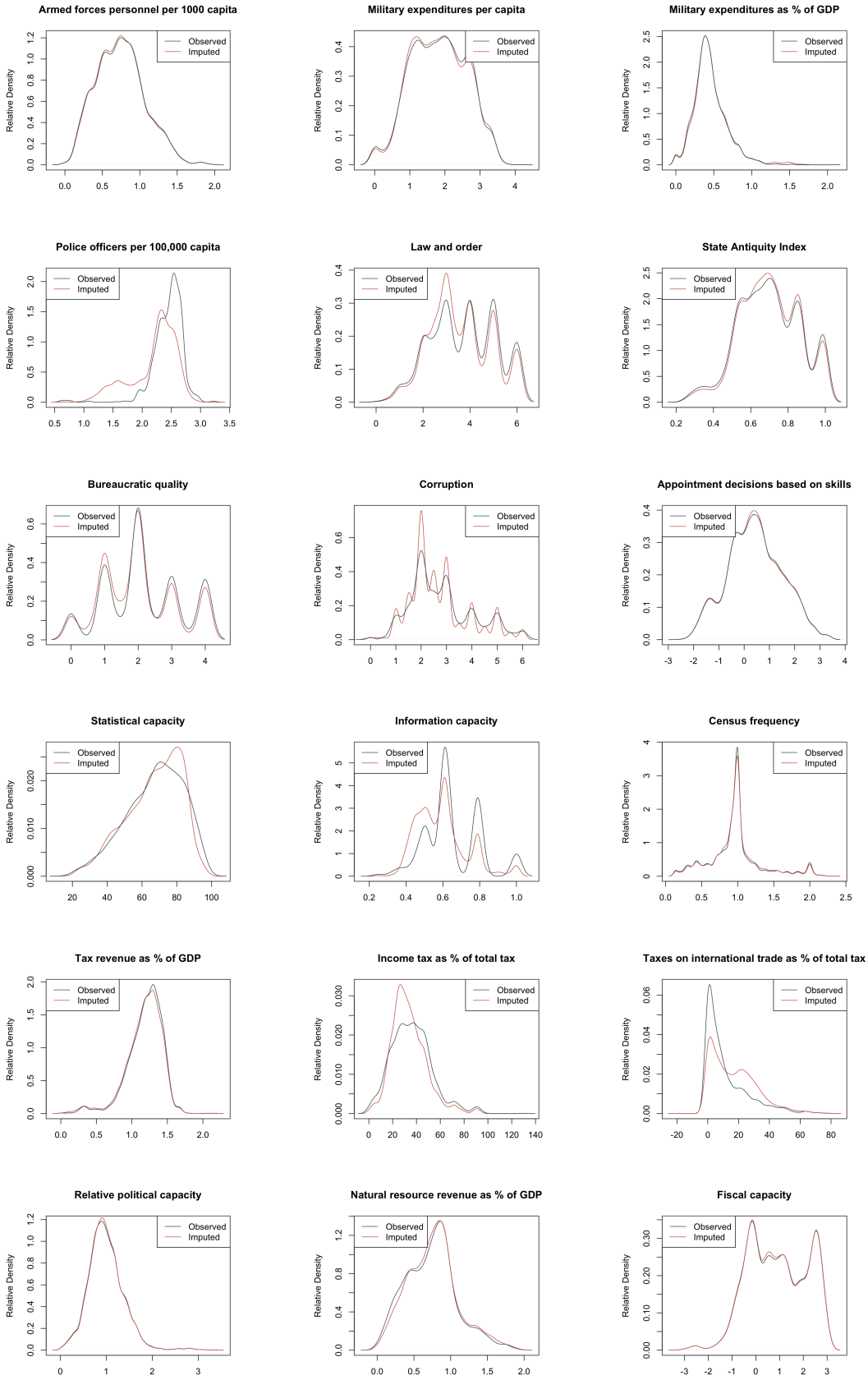
In order to avoid the inefficiencies and biases caused by the first three methods, I employ multiple imputation using R's MICE package, described more thoroughly in section 4.4. The observations are extended both in time and across space in an attempt to maximize coverage. After calculating ten imputation models through the predictive mean matching (PMM) method, I averaged the imputed values for each variable from every model and compared the empirical results using the Kolmogorov-Smirnov test, as recommended by e.g. Abayomi, Gelman, and Levy (2008). Figure 4.4 presents the Kernel density estimates for the marginal distributions of the observed data and the averaged $m = 10$ densities per variable calculated from the imputed data. It can be seen that the imputed values fit accurately within the natural range of the observed values, and that the imputed distributions generally match the observed quite well. Naturally, the largest discrepancies between the observed and imputed data can be found in the variables with the highest percentages of missing, but even these imputations should be considered to perform well.

4.3 Exploratory Factor Analysis

The theoretical perspective outlined in the foregoing chapter suggests that the underlying concept of state capacity is multidimensional. However, though distinct conceptually, there are considerable reasons to believe that these dimensions are empirically mutually constitutive and interrelated in practice (Hendrix, 2010; Hanson & Sigman, 2020). Because I expect my data to reflect one (but potentially three) underlying latent construct(s) which cannot be measured directly, I can explore the linear combinations of the observed variables (Child, 1990; Treier & Jackman, 2008). *Exploratory factor analysis* (EFA) is one approach to this, and is concerned to discover if variables form regular patterns and vary together.

When a group of variables, for whatever reason, have a great deal in common, EFA can assist in deciding whether the interrelationship can be explained by one or more latent factors. Taking as point of departure a correlation or covariance matrix,

FIGURE 4.4: Kernel density plots, observed and imputed



the method seeks to "extract" the latent variable(s) from the manifest variables by estimating the structures and pattern of relations between them (Osborne, 2014). As such, the ultimate goal is to "arrive at a parsimonious representation of the associations among measured variables" (Fabrigar, Wegener, MacCallum, & Strahan, 1999, p.275). Simply put, EFA is useful in circumstances where key variables of interest cannot be directly measured (Treier & Jackman, 2008).

The fundamental theorem of EFA is illustrated in Equation 4.1, where R_{mXm}

$$R_{mXm} - U^2_{mXm} = F_{mXp}F'_{pXm} \tag{4.1}$$

denotes the correlation matrix, U^2_{mXm} the diagnoal matrix of unique variance in each variable, and F_{mXp} the common factor loadings (Yong & Pearce, 2013). Whereas the left-hand side represents correlations and uniqueness, the right-hand side can be solved by determining the eigenvectors (factors) of the matrix with corresponding eigenvalues quantifying the proportion of common variance explained by the factor (Hendrix, 2010). In essence, the equation describes which variable is a linear combination of which common factor (Yong & Pearce, 2013, p.82).

A rule of thumb proposed by Kaiser (1960) is that one should retain factors with eigenvalues >1: these factors should be significant, and explain an important amount of variability. The *eigenvalue* is the variance of the factor, the *difference* illustrates the difference between the current and following eigenvalue, and the *proportion* gives the proportion of variance accounted for by the factor. The *factor loadings* are the regression coefficients in the multiple regression equation regressing the observed variables on the latent factors, and can be interpreted as the bivariate correlation between the variable and the factor (Hendrix, 2010, p.280). They are analyzed to identify which, and how much, variables load onto the obtained factors. Lastly, the uniqueness gives the proportion of the variance of the variable not associated with any of the latent factors, and is equal to 1 - communality.

Although I have critically assessed state capacity's *theorized* underlying latent construct and argued in favor of an aggregate measure, there is no universal agreement regarding *how* this aggregate measure should be constructed (Treier &

Jackman, 2008). EFA will assist with this through examining whether the 20 operationalizations in reality capture the hypothesized latent construct, which would provide support for an aggregate measure consisting of clustered characteristics that usefully could be described as state capacity. This is essential, given that the main purpose of the previous chapter was to demonstrate that extant disaggregated measures oftentimes are complementary, poorly conceptualized, and not seldom employed without taking construct validity into account.

Prior to running the factor analysis, I conducted two tests to confirm the suitability of my data for structure detection (Yong & Pearce, 2013). Both the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's test of sphericity yield results confirming that my data has patterned relationships, indicating that a factor analysis may be useful - with $KMO = .84$ and $p-value = .000$. Next, according to Fabrigar et al. (1999), the implementation of EFA requires the researcher to address at least five types of methodological decisions. The first concerns the study design and what measured variables to include. Naturally, within the context of EFA, the choice of variables and a theoretically sound base matter substantially for the identification of 'true' common factors. This was dealt with thoroughly in chapter 2, in which I defined the domain of interest, conceptualized the latent construct, and specified guidelines for the selection of measured variables.

The second is to determine whether EFA is appropriate. The choice of the exploratory approach is explained and justified in this particular section. When the goal of the analysis is to identify latent constructs underlying measured variables, EFA is a more sensible approach than for instance principal component analysis (PCA), where the goal is to account for variance (as opposed to correlations) in the measured variables. Furthermore, because my design is exploratory in nature (i.e., I want to explore the data structure to see whether the indicators load on three factors based on the conceptually distinct dimensions, or on one, latent variable in which the theorized dimensions instead are interrelated), EFA is preferable over the confirmatory (CFA) approach.

The third decision concerns the choice of a model-fitting, or factor-extracting, procedure. Assuming normality, Fabrigar et al. (1999, p.277) argue that maximum

likelihood (ML) is the best choice because it "allows for the computation of a wide range of indexes of the goodness of fit of the model [and] permits statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals for these parameters." However, in my case, normality tests did not indicate a significant normality. Since the assumption of multivariate normality is "severely violated", one of the principal factor methods is preferable. I chose the iterative principal factoring over the default principal factoring because the iterative procedure re-estimates communalities based on the factor loadings, and repeats the process until convergence (Osborne, 2014). Although only ML provides the opportunity to significantly test the model fit and compute confidence intervals, the realistic goal in EFA is to "obtain a parsinomious solution that provides a good approximation to the real world" (Fabrigar et al., 1999, p.280). Thus, the hypothesis of a perfect model fit is not generally of empirical interest.

Fourth, one has to determine the number of factors. I mainly rely on theoretical considerations when selecting the number of factors (3), but the Kaiser criterion, scree test and parallel analysis are also taken into consideration (de Winter & Dodou, 2012). However, these tests suggest a large number of factors which, in terms of this particular analysis, is ineffective and exceed any theoretical consensus. Nevertheless, the "correct" number of factors can never be defined indisputably. The scree plot and parallel analysis can be found in Appendix Figure C.1 and Appendix Figure C.2, respectively.

Finally, the factor analyst has to decide on the factor rotation. Factors are rotated for easier interpretation, and the goal is to "attain an optimal simple structure which attempts to have each variable load on as few factors as possible, but maximizes the number of high loadings on each variable" (Yong & Pearce, 2013, p.84). The most fundamental distinction is between the orthogonal and oblique rotations (Fabrigar et al., 1999). Because there is substantial theoretical and empirical basis for expecting my dimensions of constructs to be correlated with one another, I use the promax oblique rotation method. The orthogonal (e.g. varimax) rotation constrain factors to be uncorrelated, which would not provide a realistic representation of how factors are likely to be related. The results from the factor

analysis are presented in the next chapter.

4.4 Methodological Challenges and Limitations

Questions about the relative merits of alternative research designs and strategies pervade the social sciences (Poteete, Janssen, & Ostrom, 2010). However, instead of discrediting a certain method or its application entirely, it might be more useful to be transparent about challenges, to be aware of deficiencies, and to encourage further research (King et al., 1994; Poteete et al., 2010). As with every research project, there are multiple methodological challenges linked to thesis worth elaborating on. One has already been accounted for, namely the general difficulty of conceptualizing and measuring a latent construct. In this section, I delineate three additional challenges related to 1) the data collection process, 2) simultaneity bias and endogeneity, and 3) missingness and multiple imputation, in an attempt to best handle this issues in the analysis.

4.4.1 Data Collection

The most fundamental task when preparing data for analysis has been to select and extract indicators from multiple sources and datasets, matching certain conditions. The selected measures are, to the best of my judgement, the most conceptually fitting and frequently employed measures related to the three identified dimensions of state capacity available. During the data collection process, indicators were selected from acknowledged and generally well-trusted data sources, and this was done as carefully and consistent as possible. Although unlikely, it should be recognized that there is a possibility that errors have occurred during the collection procedure. Furthermore, as mentioned in the foregoing section, reliable factor analysis results depend on whether or not the included variables are theoretically sound and relevant to the domain of interest (Fabrigar et al., 1999). If irrelevant or poorly fitted measures are included, then factors might be obscured. This has been attempted dealt with by developing exhaustive theoretical guidelines, a careful validity assessment,

and testing the fitness of the indicators for factor analysis. However, the "omitted variable bias" could be of relevance if I have missed variables that are obviously relevant to one of the three dimensions.

4.4.2 Simultaneity Bias and Inherent Endogeneity

Simultaneity bias refers to the possibility that not only x has an impact on y, but that y at the same time effects x. As noted by for example Sobek (2010), the process of building modern, capable states involves activities to detect, eliminate or neutralize internal rivals - and more capable states are expected to be better suited to carry out these tasks. Increased state capacity should thus significantly decrease civil conflict likelihood. At the same time, it is plausible that the occurrence of civil conflict reduces state capacity (Thies, 2010). Furthermore, previous armed conflicts may affect both state capacity and the probability of a new conflict (Gates et al., 2016). Therefore, it might be that the countries at the lower end of the state capacity scale are lacking state structures *because* they are embroiled in conflict, and not necessarily the other way around. Lastly, as noted by Hendrix (2010, p.283), "all good things go together". Therefore, the dimensions and included indicators are expected to be highly collinear and endogenous themselves, but this should not cause problems for the factor analysis because an important EFA assumption is that there should be some degree of collinearity among the variables.

4.4.3 Missingness and Multiple Imputation

As accounted for in section 4.2, missing data is common in the literature on state capacity, and this thesis is no exception. Indicators considered to be important for state capacity such as bureaucratic quality and tax collection are in many countries scarce or non-existent, and this is especially true for the countries in which capacity is presumably low. All variables used in the analysis except for two contain missing values from the original data source. However, conducting EFA with missing data is problematic, which makes missingness a fundamental challenge for my analysis.

As recommended by numerous scholars for the reasons already outlined, I have

avoided listwise- and pairwise deletion (e.g. Rubin, 1976; King, Honaker, Joseph, & Scheve, 2001; Honaker & King, 2010; Zhang, 2016). In short, these methods would produce biased results unconditioned on anything else in the data unless data is missing completely at random (MCAR), which is rarely the case in quantitative studies of this kind. Additionally, these approaches would lead to an extreme loss of information, considering my data only contain 115 sets of complete country-year observations. I also refrain from using methods that create a single replacement value for each missing entry, as this result in reduced variability, shrinks standard errors, and compromise correlations. Instead, because I seek to maximize coverage in a way that reflects uncertainty, allows non-normal distributions, and replaces missing data with realistic values, I use the predictive mean matching (PMM) method in R's multiple imputations by chained equations (MICE) package (Little, 1988; Azur, Stuart, Frangakis, & Leaf, 2011; Vink, Frank, Pannekoek, & Van Buuren, 2014).

This multiple imputation model removes the under-coverage and overconfidence that results from a single imputation model by 'filling the gaps' using a predictive model that includes all available information in the observed data. For each missing value, the method forms a set of candidate donors from all similar, complete cases that have predicted values closest to the predicted value for the missing (Rubin, 1976; Van Buuren, 2018; Morris, White, & Royston, 2014). It subsequently constructs multiple completed datasets in which the imputed values vary depending on the estimated uncertainty in predicting each missing value (Honaker & King, 2010). The expected value for any missing value is the mean of the imputed values across the datasets, incorporating into the standard errors the variation across the estimates from each completed dataset.

After imputing the mean, the 'place holder' mean imputations are set back to missing, and the observed values of variable X are regressed on the other variables, operating under the same assumptions that would be the case when performing linear, logistic, or Poisson regression models (Azur et al., 2011). For example, if variable X_1 has missing values, it will be regressed on variable X_2 to X_k based on the predictive mean-defined match, whereby each variable is modeled according to its distribution. Then, the missing values for X_1 are replaced with imputations from

the regression model, and can subsequently be used as an independent variable in the regression models for the other variables (ibid.). These steps are repeated for a number of cycles, with the imputations being updated at each cycle.

In sum, PMM is an example of a "hot deck method", where missing values are imputed using values from complete cases that are in some sense close (Van Buuren, 2018). It has some distinct advantages: PMM does not assume normality, only eligible values of the missing variable are imputed, and it is less sensitive to model misspecifications because the predictive mean is only used to define a match (Little, 1988). Because PMM-imputed values in essence are real values that are "borrowed" from complete cases, the obtained imputations are more similar to expected values than imputations based on linear regression models and normal distributions (Vink et al., 2014). For instance, logical bounds for minimum and maximum values are uphold, thereby ensuring that corruption, for example, does not have values below 0 or above 6. For missing outlier values in severly skewed data, however, the method is likely to struggle because there may be no close donor values.

The method has two assumptions. First, it works best with large samples and requires a sufficiently large donor pool in order to yield acceptable inferences. In my case, the data was imputed with $m = 10$ imputations (datasets) for 50 iterations, with the default number of donors $d = 5$. Setting $d = 1$ is generally considered too low, as PMM may reselect the same donor over and over again, while setting $d = 10$ may introduce bias since the likelihood of unfitted matches increases. Setting $d = 5$ represents a compromise (Van Buuren, 2018).

Second, because the model conditions on the observed data when drawing values for the missing cells, it assumes that the missingness mechanism is at least MAR, meaning that the probability that a value is missing depends only on observed values D^{obs}, and not on the unobserved data D^{miss} (Azur et al., 2011). For example, whether a country-year's bureaucratic quality is missing should not depend on the unobserved level of bureaucratic quality. By definition, the information needed to test such an assumption is unavailable. Yet, one can scrutinize the indirect evidence by comparing the empirical distribution of the observed and imputed data. Thus, the distribution of the complete data as a whole can indicate whether the imputed

values are plausible, which is done by running imputation diagnostics. Imputations that significantly diverge from the observed data might indicate that the imputation model did not perform well, that missing data are not MAR, or both.

A first look at the missing data suggests that it might not be MAR. For instance, nine of the 20 indicators have completely missing on South Sudan and Maldives; eight have completely missing on Bhutan; seven have completely missing on Chad, Laos and Eritrea, and five have completely missing on Afghanistan, Central African Republic and Kyrgyzstan. In comparison, the maximum number of indicators that have completely missing on any Scandinavian country is two, and in East Asia, one (except for North Korea that was omitted due to high missingness). On the other hand, the countries without coverage on e.g. bureaucratic quality, corruption and law and order seem to be random, as countries usually lacking data, like Syria, Yemen, Somalia and DRC, are included. Likewise, information capacity lacks coverage on e.g. Mali and Nigeria, but also on Finland and Luxembourg.

Generally, it seems that missing is more prevalent for countries that are typically categorized as closed autocracies rather than for countries commonly deemed as "weak". One way to indirectly check the assumption is by comparing the empirical distribution of the observed and imputed data using the Kolmogorov-Smirnov (KS) test, as recommended by Abayomi et al. (2008). When applied to the variables, the tests indicate that the difference between the two samples is not significant enough to say that they have different distributions. The kernel density plot in Figure 4.4 is in a sense a visual representation of the KS test.

After considering 10 multiple imputations, a single estimate of missing values was obtained by averaging all imputed values in each dataset (Dray & Josse, 2015). When missing values are imputed, it is natural to question if and how the estimates are reliable and representative. Hence, numerous scholars have attempted to develop methods and guidelines for best practises in further analyses. Estimates are generally more reliable when datasets are larger and correlations between variables higher, which is good in this particular thesis. However, because the values obtained by multiple imputation were averaged prior to running the factor analysis, information about uncertainty around estimates is lost. This is because the aver-

aging procedure ignores the between-imputation variability. As such, the analysis is unable to differentiate which data are observed and which are missing, underestimating the uncertainty of the parameters and potentially enlargening correlations (Nassiri et al., 2018). [11]

Another challenge of using MI prior to EFA is concerned with the selection of factors to be obtained, as there is no guarantee that different methods of determining numbers of factors would propose the same decision for all the sets of imputed data (ibid.). However, as I base the factor selecting procedure mainly on theoretical grounds, this should not be the biggest issue. Although not a perfect solution, the kernel density plots yield results that the averaged imputed estimates are adequate and satisfactory for this particular study. Furthermore, the rank orderings and scatterplots presented in the next chapter look realistic. After imputation, averaging, and factoring, I checked whether the imputations and predictions seemed reasonable in comparison to other countries with similar values on variables like population size, GDP per capita and infant mortality rate. Dealing with and combining the results of different imputed datasets for use in factor analysis remains challenging, and further research on how this can be done more satisfactory and efficiently is encouraged (Dray & Josse, 2015; Nassiri et al., 2018).

The total consequence of the chosen MI procedures for the analysis is difficult to assess, but considering the fact that some variables contain a relatively high % of missing and thus imputed values, one will have to treat any results cautiously. Keeping in mind that the included variables already are *indirect* proxies of state capacity, missing values in these variables increase the uncertainty. Although using imputed (unobserved) values to predict the latent construct of state capacity is not optimal, neither is simply removing missing values. Such an approach would severely limit the analysis and seriously bias estimates. Hence, acknowledging that the undertaken methods are not perfect, they are nevertheless explained and justified, and represent one of the first steps towards bridging an important gap in the

[11] The mean and standard errors for all variables in each imputed dataset can be found in Appendix Table B.1, and a comparison of the correlation matrices before and after imputation can be found in Appendix Figure B.2.

state capacity literature. Table 4.2 presents descriptive statistics for the data that are used in the analysis.

TABLE 4.2: Descriptive statistics after log-transforming and MI

Variables	*Mean*	*Std.Dev.*	*Min*	*Max*	*Obs*	*Countries*
(log) Armed Forces Personnel per 1000 capita	.741273	.3213478	0	1.956115	4,722	165
(log) Military Expenditures as % of GDP	.4685016	.2357855	0	2.073168	4,722	165
(log) Military Expenditures per capita	1.822697	.7772521	0	4.091016	4,722	165
(log) Police Officers per 100,000 capita	2.202596	.4001865	.6180481	3.243333	4,722	165
State Authority over Territory	91.23599	10.94943	33.6	100	4,722	165
Law and Order	3.596203	1.288026	0	6	4,722	165
State Antiquity Index	.7030487	.1580526	.24	1	4,722	165
Bureaucratic Quality	2.013342	1.083543	0	4	4,722	165
Corruption	2.691705	1.180532	0	6	4,722	165
Rigorous and Impartial Public Administration	.3745235	1.497066	-3.75	4	4,722	165
Admin. Appointment Decisions Based on Skills	.42904	1.073649	-2.44	3.26	4,722	165
Statistical Capacity	66.95751	15.77055	16.66667	98.8889	4,722	165
Information Capacity	.5909524	.1354144	.24	1	4,722	165
Census Frequency	.9720372	.3643455	.1408451	2.325581	4,722	165
(log) Tax Revenue as % of GDP	1.169184	.2587788	.0000543	2.176911	4,722	165
Taxes on Income as % of Taxes	33.67114	14.72134	0	130.5398	4,722	165
Taxes on International Trade as % of Taxes	16.92104	14.75135	-18.49618	79.12088	4,722	165
Relative Political Capacity	.9995388	.3900674	.0119366	3.37442	4,722	165
(log) Natural Resource Revenue as % of GDP	.7608609	.3459463	-.0391333	1.927081	4,722	165
Fiscal Capacity	.9088941	1.215484	-3.04	2.93	4,722	165

4.5 Summary

This chapter has outlined the methods used to investigate state capacity empirically, accounted for methodological choices and challenges encountered along the way, and presented the data that will be used in the factor analysis. It has emphasized that the main challenges to this thesis are simultaneity bias and endogeneity more generally, as well as using multiple imputation to estimate an already uncertain construct. In the next chapter, exploratory principal factor analysis is employed with the aim to discover whether indicators form regular patterns and vary together.

Chapter 5

A New Measure of State Capacity

The theoretical expectation in this thesis is that although conceptually distinct, each of the three identified dimensions of state capacity exert influence on the others and are ultimately difficult to separate empirically. Although states differ in which dimension of capacities that are most strongly developed, significant strength in one should require at least some strength in the others. Thus far, I have theoretically demonstrated the advantages of an aggregate state capacity approach capturing only the core state functions that are minimally necessary for the modern state to implement desired policies. This chapter employs exploratory principal factor analysis in order to test this assumption. The results provide strong support for the proposition that the theorized dimensions are interrelated, also in practice. Therefore, a new aggregate measure is introduced.

5.1 Estimation and Results

Table 5.1 reports the results of the factor analysis. Positive factors are reported, and factor loadings are presented for the statistically significant factors with eigenvalue ≥ 1. The factor loadings for the oblique rotation represents both how the indicators are weighted for each factor but also the correlation between the indicators and the factor. Blanks represent factor loadings <.3. As the results suggests, the dimensionality of the presumed underlying concept is low, and the explanatory

TABLE 5.1: Principal factor analysis

Factor	*Eigenvalue*	*Difference*	*Proportion*	*Cumulative*
1	7.38667	4.85675	0.6693	0.6693
2	2.52992	1.41009	0.2292	0.8985
3	1.11983	0.46841	0.1015	1.0000
4	0.65142	0.06216	0.0590	1.0590
5	0.58926	0.29715	0.0534	1.1124
6	0.29211	0.14232	0.0265	1.1389
7	0.14978	0.02066	0.0136	1.1525
8	0.12912	0.05127	0.0117	1.1642
9	0.07785	0.05716	0.0071	1.1712
10	0.02068	0.04832	0.0019	1.1731

LR test: independent vs. saturated: $chi^2(190) = 6.9x10^4; Prob>chi^2 = 0.0000$

N = 4,722

	Factor Loadings			
Variable	*Factor 1*	*Factor 2*	*Factor 3*	*Uniqueness*
Bureaucratic Quality	0.8640			0.1930
Rigorous and Impartial Public Administration	0.8403			0.2461
Admin. Appointment Decisions Based on Skills	0.8192			0.3295
Corruption	0.8093			0.3959
Law and Order	0.6604			0.4403
Fiscal Capacity	0.6192			0.3581
Information Capacity	0.5881			0.4591
Taxes on Income as % of Total Taxes	0.5727			0.7418
Taxes on International Trade as % of Taxes	-0.5040	-0.3485		0.4617
Census Frequency	0.4972			0.6591
State Antiquity Index	0.4062			0.8010
State Authority over Territory	0.3428	0.3401		0.6582
Statistical Capacity		0.6964		0.3505
Relative Political Capacity		0.5947		0.6693
(log) Tax Revenue as % of GDP	0.3005	0.5851		0.3993
(log) Police Officers per 100,000 capita		0.5362	0.3580	0.5159
(log) Military Expenditures per capita	0.5163		0.8029	0.0409
(log) Military Expenditures as % of GDP			0.7777	0.3193
(log) Armed Forces Personnel per 1000 capita			0.7491	0.4087
(log) Natural Resource Revenue as % of GDP			0.6728	0.5159

Rotated factor loadings (pattern matrix) and unique variances sorted

Rotation: oblique promax

Blanks represent abs(loading)>.3

power is high: out of the three factors, factor 1 explains cumulatively 66.9% of the variance in the matrix and yields results that are consistent with the theoretical perspective that coercive, bureaucratic/administrative and extractive capacity are interrelated in practice. The only factor to represent a distinct state capacity dimension seems to be factor 3, on which the indicators representing coercive capacity and natural resource revenue load heavily.

The first factor is highly correlated with both bureaucratic-administrative features like bureaucratic quality ($r = .86$), impartial public administration ($r = .84$), corruption ($r = .80$), information capacity ($r = .58$) and census frequency ($r = .49$); extractive indicators such as taxes on income ($r = .57$) and trade ($r = -.50$), fiscal capacity ($r = .61$) and somewhat less with tax revenue as % of GDP ($r = .30$); and lastly, indicators of coercive capacity through e.g. law and order ($r = .66$), military expenditures per capita ($r = .51$), the historical presence of state institutions ($r = .40$) and state authority over territory ($r = .34$). All in all, this factor seems to capture the bundle of qualities that were hypothesized to make up the state's ability to carry out its core functions: it reflects bureaucratic, technologically advanced countries that are capable of extracting information and taxes that require higher levels of societal penetration, and ensure compliance with the law and internal order.

Table 5.2 ranks countries according to their estimated values for the three significant factors in 2019, reporting only the top fifteen and bottom eight of the 165 countries for which scores are estimated. When scrutinizing the factor 1 rank orderings, one can see that Singapore ranks among the top-fifteen countries. This is reassuring, as it likely confirms that the factor does not capture concepts more closely related to regime type and democratic governance. The complete list of 2019 rank orderings can be found in Appendix Table C.1. Here, it can be seen that United Arab Emirates and Kazakhstan rank 33 and 46 respectively, whereas *de jure* more democratic countries such as South Africa and Albania rank 62 and 68.

Interestingly, the measure capturing natural resource revenue as % of GDP is not significantly correlated with the first factor in either direction ($r = .17$) despite the literature linking higher proportions of primary commodity revenue to lower levels of extractive and bureaucratic-administrative capacity. Neither are the indic-

TABLE 5.2: Principal factor analysis rank orderings, 2019

Rank orderings	*Factor 1*	*Factor 2*	*Factor 3*
1	Denmark	Serbia	*Libya*
2	Germany	Croatia	Saudi Arabia
3	Luxembourg	Bulgaria	*Syria*
4	Sweden	Kyrgyzstan	Kuwait
5	New Zealand	Belarus	Oman
6	Ireland	Mongolia	Bahrain
7	United Kingdom	North Macedonia	*Yemen*
8	Norway	Armenia	*Russia*
9	Switzerland	Moldova	*Israel*
10	Netherlands	Montenegro	Qatar
11	Australia	Poland	*Somalia*
12	Singapore	El Salvador	Azerbaijan
13	Finland	Slovenia	Algeria
14	Canada	Georgia	Jordan
15	Belgium	Mauritius	Lebanon
. . .			
158	*Congo, DR*	UAE	Liberia
159	*CAR*	*South Sudan*	Benin
160	*Syria*	*Syria*	*Mozambique*
161	*Haiti*	*Libya*	Madagascar
162	*Libya*	*Iraq*	Malawi
163	*South Sudan*	*CAR*	Haiti
164	*Yemen*	*Yemen*	Panama
165	*Somalia*	*Somalia*	Costa Rica

Italics denote that the country experienced a civil conflict in 2019.

ators related to military ($r = -.16$) and police ($r = .05$) personnel significantly correlated to factor 1. The former can probably be explained by the fact that factor 1 correlates with more direct measures of both extractive and bureaucratic capacity. As Hendrix (2010, p.282) explains, "taxation is key because it is the incentives (or rather, the lack thereof) to access the resources of the governed that motivate our most compelling theoretical accounts of the development of the modern state".

When it comes to the measures capturing coercive personnel, especially two

explanations are plausible. First of all, it is likely that factor 1 misses the measures linked to state employment of security personnel because of the endogeneity problem already accounted for. States are expected to expand their security forces when threatened by internal disorder, thereby weakening the relationship. This also seems like a reasonable explanation for why military expenditures as % of GDP is not significantly correlated with the factor ($r = -.19$). Second of all, because other measures of coercive capacity are moderately to strongly correlated with factor 1, it might be that it is not the number of armed forces personnel that matter for the level of coercive capacity, but rather, as suggested by Hanson and Sigman (2020), the administrative organization and technological sophistication.

Accordingly, I interpret the results from factor 1 as consistent with the theoretical argument made that it is more constructive to view the extractive, coercive, and administrative dimensions of state capacity as mutually interrelated and codependent, not only in theory, but also in practice. An administratively sophisticated coercive apparatus is needed to contain internal threats and to ensure order and compliance with the law, and in order to fulfill this task, the state requires a certain level of extracted revenue. This, in turn, requires the bureaucratic-administrative capacities to reach populations and to collect and manage information. The results are furthermore consistent with recent contributions (Berwick & Christia, 2018; Hanson & Sigman, 2020). I will term this factor State Capacity, and the new variable will be tested in different validity and utility checks in the next chapter.

Factor 2 seems to capture some kind of bureaucratic "stateness" that is more coercive. The indicators that load most heavily on this factor are statistical capacity ($r = .69$), relative political capacity ($r = .59$), tax revenue ($r = .58$), taxes on trade ($r = -.34$), police personnel ($r = .51$) and state authority over territory ($r = .34$). According to the 2019 rank ordering, factor 2 is heavily dominated by former Soviet states, in which state evasion is avoided through the semi-voluntary exchange of state enforcement services for tax revenues. In a sense, this pattern can be linked to Olson's (1993) theory on stationary banditry, Bates' (2008) "specialists in violence" and Tilly's (1985) "racketeering". In short, the provision of security of person and property is uphold by coercive control (or the threat thereof) in exchange for citizens'

taxes. Yet, ironically, the ability to credibly threaten violence does not necessarily encourage development, nor the incentives to invest in long-term capabilities that would ease the dependency of coercion. Although the high-ranking states in factor 2 are not embroiled in conflict, nor are they recognized for being the most "capable" across state arenas beyond the coercive.

Factor 3 on the other hand seems to capture the effects of endogeneity mentioned in relation to factor 1, and is heavily influenced by military expenditures per capita ($r = .80$), military expenditures as % of GDP ($r = .77$), armed forces ($r = .74$) and police ($r = .35$) personnel, as well as natural resource revenue as % of GDP ($r = .67$). The 2019 rank orderings confirms this, as all the top-ranking countries reflect either state dependence on oil, gas, and/or mining commodities, or prevailing civil conflict (or both). As mentioned in chapter 3 section 3.2, there exists widespread uncertainty about the true causal mechanism linking measures of military aspects to coercive capacity. Thus, to a large extent, factor 3 demonstrates the flaws of using these measures as proxies for state capacity. A large military force is instead, at least in this analysis, a symptom of increased threat perceptions, leading the state to compensate through investments in repression and defense. The "least capable" state according to factor 3 is Costa Rica, which has no standing armies and only limited mineral resources. Moreover, factor 3 fails to establish a relationship with extractive features like tax revenue ($r = -.14$) and all bureaucratic-administrative aspects. In sum, factor 3 is bounded at the top by states in which the monarch is the direct beneficiary of the country's natural resource wealth (e.g. Saudi Arabia, Kuwait, Oman and Bahrain), and noticeably portray neopatrimonialism and extensive military- and natural resource-presence.

5.2 Exploring State Capacity

Table 5.3 presents the descriptive statistics for the new aggregated variable from factor 1, State Capacity. It is scaled from -2.448843 (Somalia, 1993) to 2.382562 (Denmark, 2012), with a mean and standard deviation of -0,000000000131 and .9755503, respectively. It covers all the countries in the dataset (165) from 1991 (or from the

first year or independence) to 2019 with a total of 4,722 observations, offering much greater coverage than comparable indices commonly employed for research in the post-Cold War era.

TABLE 5.3: State Capacity descriptive statistics

	Obs	*Mean*	*Std.Dev.*	*Min*	*Max*
State Capacity	4,722	-1.31e-10	.9755503	-2.448843	2.382562

Furthermore, because it is constructed based on a thorough reconceptualization emphasizing the core dimensions of the state, the measure is expected to capture the essence of the concept more comprehensively and satisfactory in terms of validity than previous contributions that is focused on a single indicator or dimension. Additionally, it should have broad utility and be of relevance to multiple fields of research. It would for instance be interesting to see how the State Capacity measure performs in predicting the kinds of societal outcomes that are widely associated with State Capacity, like distribution of public services, health- and schooling, drug control and crime rates, and the size and magnitude of (illegal) economic activity. However, due to the scope limitations of this thesis, I leave these tests to further research, and will in the next chapter instead focus on the measure's general validity and predictive performance when it comes to conflict outbreak.

5.3 Summary

Taking as point of departure a thoroughly explained reconceptualization of state capacity, this chapter has explored the interrelationship between 20 carefully selected indicators relating to its three identified dimensions, using principal factor analysis. The results are consistent with the pervasive theoretical assumption: the underlying dimensionality of state capacity is low, leading one factor to explain close to 67% of the variance in the measures. This factor, State Capacity, draws on indicators representing all the dimensions that were hypothesized to make up the state's ability to carry out its core functions. In the next chapter, the new measure is examined in terms of its face validity, content validity, convergent validity and predictive validity.

Chapter 6

Diagnostics and Application

The next step is to investigate whether the new State Capacity measure constructed from factor 1 in the foregoing chapter behaves in the expected manner. Therefore, in this chapter I attempt to fully validate the new variable by running different types of validity and utility checks, as recommended by Adcock and Collier (2001) and Seawright and Collier (2014). The new measure is examined in terms of its face validity, content validity, convergent validity and predictive validity. If State Capacity truly captures states' abilities to implement policies and execute core functions, validity tests should confirm that it is empirically related to other attempts to measure the concept, that it is different from measures that might be correlated with but nevertheless are conceptually distinct from State Capacity, and that it preforms well in predictive tests. In terms of predictive power, I examine State Capacity's effect on intrastate conflict onset likelihood using logistic regressions. If the tests yield positive results, the measure should be considered validated.

6.1 Face Validity

Face and content validity can be tested by investigating whether "on its face" the measure makes sense and seems like a good translation of the construct (Seawright & Collier, 2014). When examining the full list of 2019 rank orderings in Appendix Table C.1, one notices that the higher-scoring countries altogether correspond to

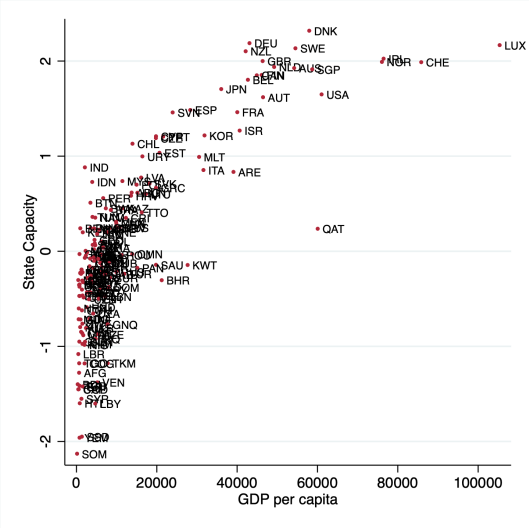
the countries one may expect to have higher capacity. Among the top 35, one can identify Scandinavian countries like Denmark, Sweden, Norway and Finland; other West-European states like Germany, Luxembourg, Switzerland, United Kingdom and Belgium; Latin American countries like Chile and Uruguay; East- and South-East Asian states like Singapore, Japan, South Korea and Malaysia; Middle-Eastern countries like Israel and the United Arab Emirates; and North-American countries like Canada and the United States. At the lower end, one finds countries ranging from Somalia, DRC and South Sudan to Haiti, Venezuela and Burundi. This too is plausible, as they have either recently experienced civil conflict, are currently embroiled in conflict, and/or they are fundamentally lacking abilities to carry out the core state functions captured by the measure. Of course, states differ in terms of which dimension of capacities are most strongly developed, but the interplay is assumed to be significant: convincing strength in any one dimension likely requires at least some strength in the others (Hanson & Sigman, 2020).

Although different concepts, it is reasonable to suspect State Capacity to be highly correlated with outcome variables such as economic development and infant mortality rates (IMR). Figure 6.1 plots State Capacity scores for all the countries in the dataset represented by their ISO3 country code in year 2019 with their respective levels of GDP/capita and IMR the same year. As suspected, State Capacity is quite strongly correlated in the expected direction with both: $r = .74$ for GDP/capita and $r = -.63$ for IMR. It *makes sense* that the countries with the highest State Capacity rankings also are wealthier in terms of GDP/capita, and this relationship should represent both causes and effects of state capacity and development simultaneously.

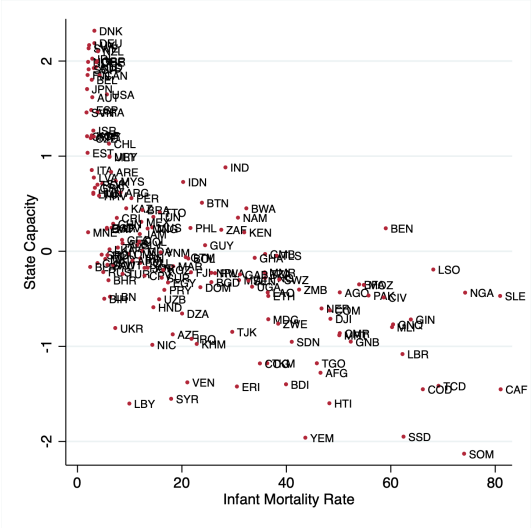
Moreover, the observed outliers fit well with what one would expect from the neopatrimonialist and oil-rich monarchies in the Middle East: Qatar, Kuwait, Bahrain and Saudi Arabia all have higher levels of GDP/capaita than other countries with similar State Capacity scores. Yet, despite their levels of income, it seems like they are not extraordinarily capable of carrying out the key functions of the state, as their State Capacity scores are lower than for instance India, Indonesia, Peru, Chile and Tunisia, all of which have lower incomes than the aforementioned states. This is a good sign, indicating that State Capacity indeed reflects aspects not captured

FIGURE 6.1: Scatterplot of state capacity, GDP/capita and IMR, 2019

(A) GDP/capita



(B) Infant mortality rate



by economic development. Although they are natural resource rich and military-wise strong countries, they are not characterized by administratively well-organized, impartial and meritocratic bureaucracies capable of implementing a wide range of policies. This is not necessarily because they are unable to develop such functions, but rather because it is not politically prioritized.

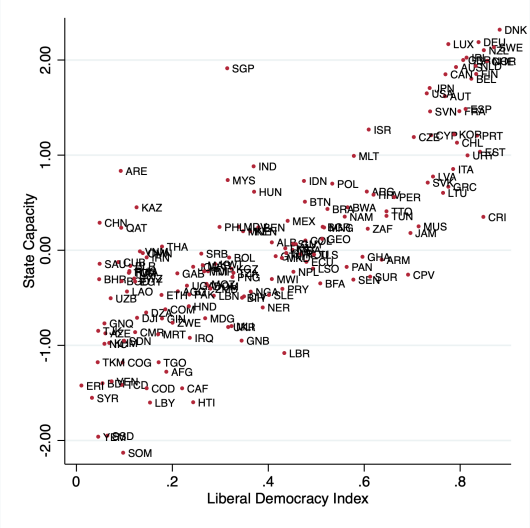
Similarly, as the countries with lowest State Capacity also are the ones regularly embroiled in conflict, one would expect IMR to be higher for these countries. Evidently, this is especially true in for instance South Sudan, DRC, Chad, Somalia and the Central African Republic, theoretically supporting the predictions of conflict recurrence and the "conflict trap" (P. Collier et al., 2003; Hegre, Strand, Gates, & Nygård, 2011; Gates et al., 2012; Thies, 2015; Gates et al., 2016). Although higher state capacity should deter intrastate armed conflict onset, these types of conflicts tend to aggravate the same structural forces that helped caused them in the first place, hampering the government's ability to invest in state capacity, to handle conflicts in a satisfactory manner, and to implement conflict-reducing policies. Thus, I interpret the empirical relationships between State Capacity and GDP/capita, and State Capacity and IMR, as "good signs" for State Capacity's face validity.

Furthermore, as thoroughly explained in chapter 3, it is important to check

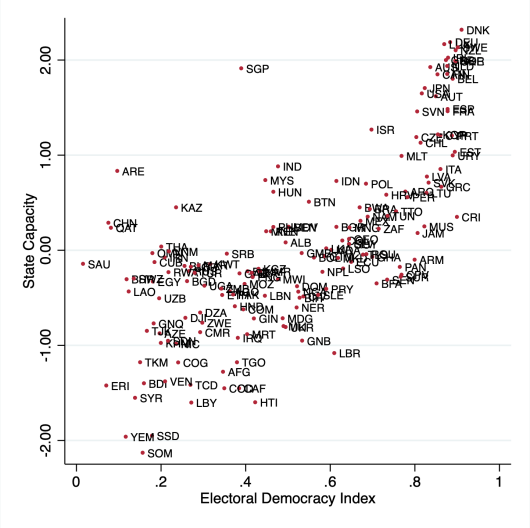
that the measure succeeds in differentiating itself from concepts more directly connected to regime type. That Singapore ranks among the top-fifteen countries is reassuring, but we are also interested in the general relationship between the commonly conflated measures. Figure 6.2 plots State Capacity scores for all countries in 2019 represented by their ISO3 country code, with V-Dem's Liberal Democracy (6.2a) and Electoral Democracy (6.2b) scores (Coppedge et al., 2021). State Capacity is strongly correlated with both: $r = .81$ for the Liberal Democracy Index and $r = .73$ for the Electoral Democracy Index.

FIGURE 6.2: Scatterplot of state capacity and regime type, 2019

(A) Liberal Democracy Index



(B) Electoral Democracy Index

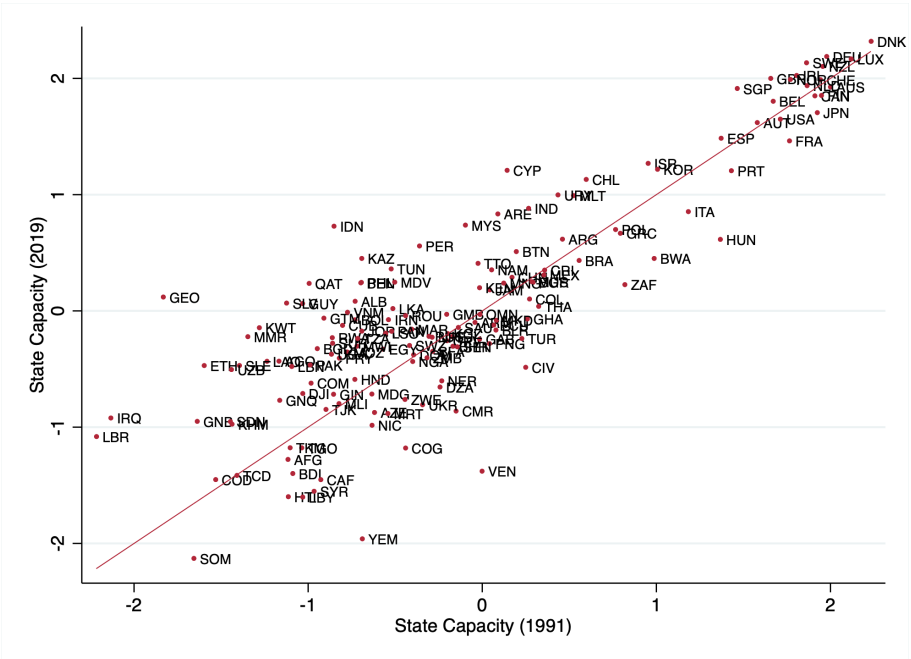


As the figures depict, more capable states tend to also be democratic. This pattern too is consistent with theory, which usually predicts that democratic countries' public administrations to a greater extent are impartial and bureaucratic and less corrupt, which is expected to improve abilities to effectively implement policies, compared to especially intermediate regimes. Less democratic and more autocratic countries can of course also be capable of implementing policies across its territories, which the figure clearly exemplifies. However, these countries *tend to* rely more heavily on coercive capacities than do democratic ones. Examples of very to fairly capable but less (liberal) democratic, or autocratic, states are Singapore, United Arab Emirates, Cuba, Israel, Malaysia and China. I interpret this as an

encouraging sign for the measure's face validity.

Finally, I test the face validity of the measure's temporal variation. There is general agreement in the literature on state formation that state capacity, despite interstate differences, should increase over time. Figure 6.3 plots State Capacity scores for all countries, with 1991 scores on the x-axis, 2019 scores on the y-axis, and a 45° line in between. The results are compatible with theory, as the relationship between the State Capacity measure in different years is significantly positive. As expected, the countries that ranked high in 1991 also have high scores in 2019. Moreover, quite a few countries have had their capacity decreased. The most apparent include Somalia, Yemen, Venezuela, Congo, and Ivory Coast. On the other hand, capacity rose most in states like Liberia, Iraq, Georgia, Indonesia, Qatar, Myanmar, Peru and Cyprus. The over-time variation for the most capacity-decreasing and capacity-increasing countries can be found in Appendix Figure D.1. These trends provide additional reassurance for the measure's face validity, and are consistent with what one would expect considering that the measure's most influential aspects include bureaucratic quality, administrative impartiality, corruption, law and order, fiscal capacity and tax extraction.

FIGURE 6.3: Scatterplot of state capacity, 1991 and 2019



In Appendix Table D.1, one can find the pairwise correlation of State Capacity with its base indicators, confirming once again that the latent construct of state capacity is strongly related to the broad set of included state functions. Unless I have overlooked other essential indicators that are both theoretically and empirically important for either of state capacity's three dimensions, this measure should thus far be satisfactory in terms of face- and content validity.

6.2 Convergent Validity

The next step of validity check I examine is the measure's convergent validity; that is, the degree to which State Capacity is similar to other aggregate measures that it theoretically should converge on (Seawright & Collier, 2014; Vaccaro, 2020). However, as noted by both Carmines and Zeller (1979) and Adcock and Collier (2001), because state capacity is inherently latent, "true" measures do not exist against which this validation can be carried out. Therefore, I have chosen to compare the State Capacity measure to other aggregate measures that attempt to capture state capacity, strength, quality, and/or fragility.

With bivariate correlations, I can assess how the variables are related as well as the strength of their relationships. This procedure is conventionally used as a tool of measurement validation, and strong correlations should be evidence of satisfactory convergent validity (Vaccaro, 2020; Hanson & Sigman, 2020). At the same time, I expect the correlations to also illustrate differences, as most other indexes include components directly connected to for instance economic development, regime type, civil liberties, or grievances. Furthermore, the fact that the measures obviously are aggregated differently should also, at least partially, explain the expected disparity.

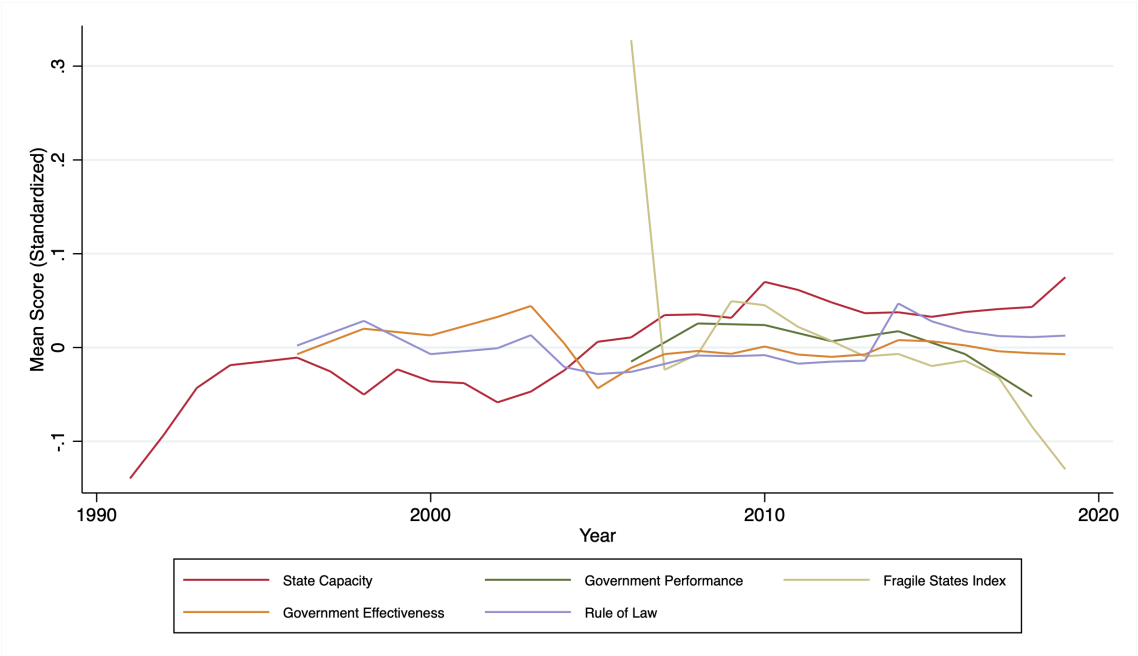
Table 6.1 confirms the assumption that State Capacity is strongly correlated in the expected direction with several similar measures. It is particularly similar to WGI's measure of Government Effectiveness ($r = .93$), Rule of Law ($r = .91$) and Regulatory Quality ($r = .89$), Fund for Peace's Fragile States Index ($r = -.87$), and Center for Systemic Peace's State Fragility Index ($r = -.79$). It is the least correlated with CPIA's measure of Quality of Public Administration ($r = .60$).

TABLE 6.1: Pairwise correlation of State Capacity with similar measures

Indicators	*Source*	*r*	*N*
Government Effectiveness	Worldwide Governance Indicators	0.93	3,442
Rule of Law	Worldwide Governance Indicators	0.91	3,443
Regulatory Quality	Worldwide Governance Indicators	0.89	3,443
Fragile States Index	Fund for Peace	-0.87	2,225
State Fragility Index	Center for Systemic Peace	-0.79	3,877
Government Performance	Bertelsmann Stiftung	0.77	867
Governance Index	Bertelsmann Stiftung	0.72	867
Political Stability and Absence of Violence	Worldwide Governance Indicators	0.69	3,443
Stateness	Bertelsmann Stiftung	0.66	867
Quality of Public Administration	Country Policy and Institutional Assessment	0.60	930

Moreover, Appendix Table D.2 shows that this pattern also is consistent for State Capacity's constitutive components, as most of the alternative measures correlate highly with all, or close to all, variables that make up State Capacity. Yet, the base indicators tend to be the most correlated with State Capacity. Figure 6.4 illustrates the over-time mean variation for State Capacity and four of the other measures when the scores are standardized.

FIGURE 6.4: Over-time variation in similar measures



A valid measure should not only be related to *similar*, constructs; it must additionally be able to adequately differentiate itself from other concepts that not necessarily reflect the same thing, but are related nevertheless. I demonstrated this with certain measures in the last section. Although State Capacity, unsurprisingly, is strongly related to e.g. GDP/capita and liberal democracy, it is reassuring that it is even more highly correlated with measures that attempt to capture the same concept. All in all, these findings indicate a high convergent validity, and the measure seems to perform well not only in differentiating itself from related yet conceptually distinct concepts, but also because it behaves in the expected manner when compared to other measures that perhaps are harder to distinguish.

6.3 Interchangeability and Predictive Validity

In this section, I demonstrate the utility of the new aggregate measure using the level of State Capacity as a predictor for intrastate armed conflict *onset*. I consider this the most challenging test for the measure's overall power and applicability, as I control for (although not exhaustive) a variety of other aspects of the state that are generally considered decisive for whether or not a country will experience conflict. First, I replicate one of the most-cited studies in the state capacity-armed conflict literature. It should be stressed that this is not intended to criticize the study, but rather to test state capacity's interchangeability and whether undisclosed flexibility can be uncovered by adding the new measure. Thereafter, I run my own logistic regression for 1992-2017, testing whether State Capacity significantly can predict conflict onset. If these tests yield robust results, one should be fairly confident that the measure is indeed capturing aspects of the state that are key for its ability to execute its core functions and thus avoid conflict onset, that other measures do not.

6.3.1 Replication

As a starting point, I replicate Fearon Laitin's (2003) logistic regression models 1 and 2 for the years 1992-1999. This is done both with and without my measure of

State Capacity. The limited time series is due to the fact that their data stop in

TABLE 6.2: Replication of Fearon & Laitin, 1992-1999

	Model			
	(1)	(2)	(3)	(4)
	Civil War	Ethnic War	Civil War	Ethnic War
State Capacity^{-1}			-1.869***	-1.973***
			(-4.07)	(-4.00)
Prior War	-0.908	-0.671	-1.336*	-1.110
	(-1.65)	(-1.15)	(-2.39)	(-1.89)
(log) Per Capita Income^{-1}	-1.112**	-1.071**	-0.418	-0.422
	(-3.00)	(-2.70)	(-1.07)	(-1.02)
(log) Population^{-1}	0.369*	0.379*	0.453*	0.486*
	(2.11)	(2.07)	(2.32)	(2.39)
(log) % Mountainous	0.136	-0.009	0.116	-0.0632
	(0.89)	(-0.05)	(0.60)	(-0.29)
Noncontiguous State	0.765	1.138	1.246	1.732*
	(1.13)	(1.60)	(1.75)	(2.26)
Oil Exporter	0.659	0.548	0.354	0.264
	(1.12)	(0.87)	(0.55)	(0.38)
New State	2.995***	2.925***	2.105**	2.047*
	(3.94)	(3.42)	(2.69)	(2.34)
Instability^{-1}	0.570	0.340	0.449	0.103
	(1.15)	(0.62)	(0.84)	(0.17)
Democracy^{-1}	-1.162	-1.488	-0.419	-0.654
	(-1.68)	(-1.80)	(-0.59)	(-0.77)
Ethnic Fractionalization	-0.154	-0.310	0.252	-0.158
	(-0.17)	(-0.30)	(0.23)	(-0.13)
Religious Fractionalization	-0.600	0.602	-0.479	0.884
	(-0.48)	(0.44)	(-0.37)	(0.61)
Constant	0.712	0.081	-6.538	-7.022
	(0.23)	(0.02)	(-1.82)	(-1.84)
N	1220	1220	1220	1220

Values in parentheses are t statistics

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1999, and because my measure of State Capacity starts in 1992 when it is lagged. In their analysis, they code civil war as 1 for all country-years in which a civil war started and 0 for all others. The main independent variable of interest in this setting, which they claim represents state capacity, is (log) per capita income. Table 6.2, model 1 and 2 reports the estimates obtained from the replicated logistic regression. In model 3 and 4, I have added my alternative State Capacity measure.

The total number of conflicts is 25, spread over 23 countries. The findings are consistent with Fearon Laitin's (2003, p.76) argument that states of lower capacity are associated with a higher risk of civil war. This is confirmed by all four models. However, as can be seen from model 3 and 4, this is not due to levels of income. Interestingly, at least on the surface, my State Capacity measure performs considerably better than the widely used GDP/capita measure in predicting both civil and ethnic war. In fact, GDP/capita even drops out out of significance when State Capacity is added, and mountainous areas have no noticeable effect. Population and the new state variables are the only ones that are significant at the .05 level across all models, meaning that higher populations and newly independent states represent significant risk factors. The State Capacity to civil war $OR = .154$, and for each additional unit decrease in state capacity, the odds of observing civil war onset increases by 84.6%.

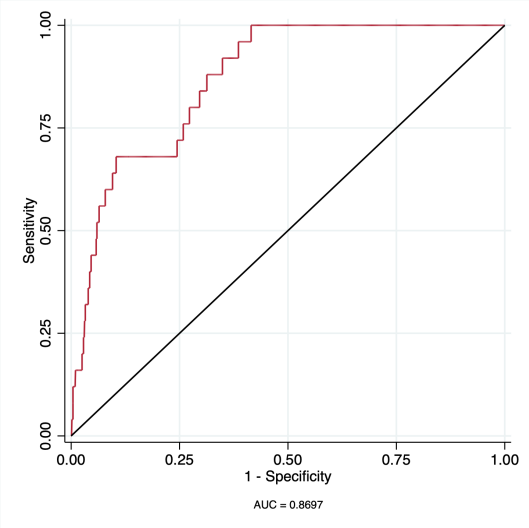
In comparison, Fearon & Laitin's measure of state capacity (per capita income) predicts that for each one-unit increase in log per capita income, the odds of civil war decrease by 67.1%. In Figure 6.5, one can find the plotted receiver operating characteristic (ROC) curves for the two compared models, displaying each model's sensitivity and one minus specificity. This is a common way to assess how well the logistic regression models fit the data, and calculates the probability that a given observation has a positive outcome, based on the values of the predictor variables. In short, sensitivity refers to the probability that the model predicts a positive outcome when indeed the outcome is positive (civil war), and specificity refers to the ability to predict a negative outcome when indeed the outcome is negative. Hence, we want models to obtain both high sensitivity and high one minus specificity, hugging the top left corner of the plot. The AUC range from 0-1, and higher comparative scores

indicate better performance in classifying outcomes correctly.

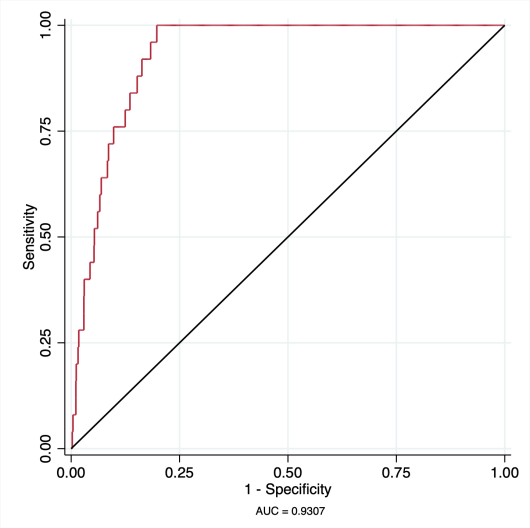
From the figure, it can be seen that both models seem to perform well in predicting the possible 25 civil wars. Still, the state capacity model is the superior, with an AUC score of .931 compared to Fearon & Laitin's per capita income model with AUC = .867. Akaike's information criterion and the Bayesian information criterion confirms this. Given two models, the one with the smaller AIC and BIC fits the data better. Both criteria pick the state capacity model as the better-fitting.

FIGURE 6.5: Receiver operating characteristic (ROC) curves

(A) Per capita income model



(B) State Capacity model



These are interesting findings. Although the results confirm Fearon Laitin's (2003) state capacity hypothesis, model 3 and 4 suggest that GDP/capita is a flexible measure and refute the argument made that "it is a proxy for a state's overall financial, administrative, police, and military capabilities" (2003, p.80). According to my predictive models, both civil and ethnic wars are significantly more likely to occur if *state capacity* (as measured by the inherent coercive, administrative and extractive abilities of the state) is lower, populations larger, and the state is newly independent. Economic development has no substantive effect when state capacity is controlled for. Of course, endogeneity issues like reverse causality and omitted variable bias are hard to unpack in these types of models, and might bias results for instance by inflating standard errors.

However, one *can* be fairly confident that 1) the GDP/capita measure is inadequate in capturing the broad range of state capacities necessary to prevent conflict, that 2) thus far, the alternative State Capacity performs better and has greater civil war predictive power than the former, and that 3) the measures are interchangable in the way that they lead to similar empirical findings, although GDP/capita drops out of significance when State Capacity is introduced. Because the causal mechanism linking income per capita to civil war is unclear, it is hard to establish exactly why the measure lost its significance when introducing State Capacity. Inded, it is exactly because the causal mechanism linking GDP/capita to civil war is unclear that I discourage using it as a proxy for state capacity. What is clear is nevertheless that they capture different aspects to the state.

6.3.2 State Capacity and Intrastate Conflict Onset

Next, I run a logistic regression analysis of my own. The dichotomous dependent variable is obtained from the UCDP Onset Dataset version 19.1, and measures the onset of intrastate armed conflicts with $\geq$ 25 battle deaths (N.-P. Gleditsch et al., 2002; Pettersson, Högbladh, & Öberg, 2019)[12]. The variable is coded as 1 if the conflict is new or if there has been more than two years since the last observation of the conflict, whereas no observation of conflict, or observations of conflicts that do not satisfy the these thresholds, correspond to the coding 0. Analysis is conducted for 1992-2017, as 2017 is the last year with observations in the onset variable and 1992 is the first year of complete observations after the variables are lagged. The total number of observed conflict onsets is 163.

The list of control variables is by no means exhaustive, and neither should it be. My aim is to test the validity and power of the State Capacity measure when it comes to predicting intrastate conflict outbreaks when a limited number of other, commonly regarded important aspects to the state are controlled for. Conducting a full, theory-driven regression analysis in which complex inter-variable relationships are analyzed is beyond the scope of this thesis, but encouraged in further analyses.

[12]Data can be downloaded from `https://ucdp.uu.se/downloads/index.html#onset`

Thus, the included control variables should be deemed adequate for the purpose of validity testing.

TABLE 6.3: The effect of State Capacity on conflict onset, 1992-2017

	Intrastate conflict onset				
	(1)	(2)	(3)	(4)	(5)
State Capacity[a]	–0.816***	–0.727***	–0.494**	–0.526**	-0.455*
	(0.221)	(0.195)	(0.204)	(0.228)	(0.238)
(log) GDP/capita[a]	–0.229	–0.181	–0.674**	–0.671**	–0.679**
	(0.327)	(0.288)	(0.335)	(0.331)	(0.331)
(log) Population[a]	1.295***	1.174***	1.145***	1.152***	1.135***
	(0.235)	(0.191)	(0.181)	(0.177)	(0.185)
Peace Years		–0.133	–0.106	–0.104	–0.108
		(0.103)	(0.100)	(0.100)	(0.100)
(log) Natural Resource Wealth[a]			0.951***	0.965***	0.941***
			(0.285)	(0.298)	(0.295)
Electoral Democracy Index[a]				0.160	
				(0.727)	
Liberal Democracy Index[a]					–0.183
					(0.855)
N	4213	4213	4213	4213	4213

Notes: (a) lagged one year; (b) dichotomous; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Values in parentheses are robust standard errors clustered by country-ID
Coefficients for constants and cubic splines are not shown

Table 6.3 presents the results from five logistic regression models, testing whether the State Capacity measure is able to predict intrastate armed conflict onset even after one by one controlling for aspects to the state that represent economic development, population size, natural resource wealth, and regime characteristics. I use Beck, Katz & Tucker's (1998) technique to address the problems of serial correlation and temporal dependence by including a counter for the number of peace years prior to each onset and three cubic splines. The cubic splines are however not presented in the table. Additionally, I lag all time-varying independent variables by one year to at least partly address the concerns related to simultaneity and endogeneity. Because employing fixed effects is not feasible, I address the issue of within-entity dependence by clustering standard errors by country. Thus, the parentheses in the

table represent cluster-robust standard errors.

In each model, State Capacity is a statistically significant predictor, and decreasing State Capacity is strongly associated with higher probabilities of intrastate conflict outbreak. In model 1, I find that each one-point decrease in State Capacity is associated with a 55.8% increase in the odds of intrastate conflict onset, controlling for GDP/capita and population size. This finding is significant at the .01 level, and economic development has no robust effect. Model 2 controls for the number of peace years prior to the conflict onset, which does not significantly reduce the predictive power of State Capacity: the variable is still significant at the .01 level, and the $OR = .48$, resulting in a 52% decreased odds for conflict outbreak for every one-unit increase in State Capacity.

FIGURE 6.6: Predicted probabilities for conflict onset, model 1 and 2

(A) Model 1 (B) Model 2

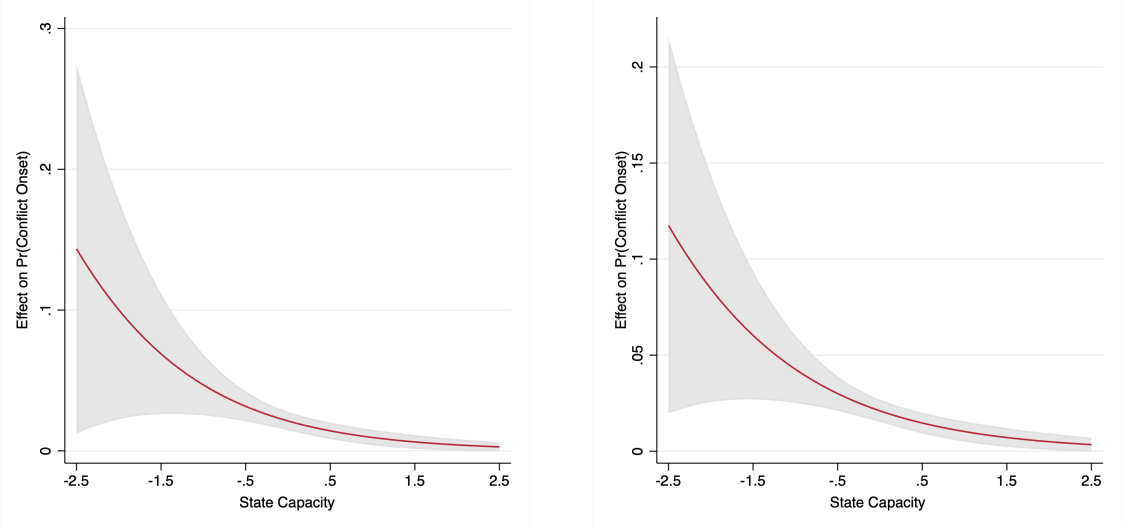


Figure 6.6 and Table 6.4 illustrate the predicted probabilities for conflict outbreak and their 95% confidence intervals when all control variables are held constant at their means. As can be seen, the predicted probabilities of conflict outbreak clearly increases with lower levels of state capacity, although confidence intervals increase with decreasing state capacity. However, that confidence intervals are large is not surprising nor necessarily a bad thing; they simply reflect the possible range around the estimate and the margins of error. In these examples, they illustrate that states

of greater capacity generally tend to have very low probabilities for conflict onset, while for states of lower capacity, variability is larger: some "low capacity states" are very likely to experience conflict outbreak, while others have predicted probabilities close to 0.

TABLE 6.4: Predicted probabilities for conflict onset, all models

	Predicted Probabilities of conflict onset				
State Capacity score	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
-2	10.03%	8.45%	5.40%	5.74%	5.00%
	[.02238, .17825]	[.02544, .14364]	[.01390, .09416]	[.00995, .10499]	[.00557, .09449]
-1	4.70%	4.27%	3.36%	3.47%	3.23%
	[.0253, .06869]	[.0251, .06031]	[.01960, .04775]	[.01889, 05064]	[.0164, .04824]
0	2.13%	2.11%	2.08%	2.08%	2.07%
	[.01477, .02792]	[.01531, .02696]	[.01518, .02645]	[.0151, .02648]	[. 01512, .02641]
1	0.95%	1.03%	1.28%	1.24%	1.32%
	[.0040, .01507]	[.00504, .01562]	[.00606, .01955]	[.00539, .01944]	[.00573, .02083]
2	0.42%	0.50%	0.78%	0.73%	0.84%
	[.00011, .00838]	[.00069, .00935]	[.00083, .01487]	[.00013, .01461]	[-.00002, .01696]

Notes: brackets are 95% confidence intervals; all other control variables are held at their means

For model 1 in the regression output, the predicted probability of observing a conflict outbreak for states with a capacity score between .4 and -.4 is 2%, compared to 12% for states scoring $\leq$ -2, given that per capita income and population size are held at their means. Similarly, with the independent variables at their means in model 2, the probability of conflict onset is 8.45% for states that score -2 on State Capacity, while it is 0.5% in a state scoring 2.

Model 3, 4 and 5 additionally control for natural resource wealth and democracy. Data on natural resource wealth is retrieved from ICTD, and calculated as the difference between total revenue and total non-resource tax revenue, reported in % of GDP. Both the Electoral and the Liberal Democracy variables are obtained from V-Dem (Coppedge et al., 2021). Model 3 and 4 illustrate that State Capacity is still significant at the .05 level, together with GDP/capita, population size, and natural resource wealth. When controlling for income, population, number of peace years, cubic splines, resource wealth and electoral democracy, the odds of conflict outbreak is 41% higher for every one-unit decrease in State Capacity.

Interestingly, the odds of conflict onset is estimated as 2.5-2.6 times greater for every one-unit increase in the log number of natural resource wealth. This is consistent with the results from factor 3 in the EFA employed in chapter 5, and provides further confidence in the proposition that measures related to natural resource wealth and armed forces are poor proxies for coercive capacity. In *preventing* conflict onset, the results suggests that bureaucratic-administrative organization in all three dimensions are decisive determinants: this type of capacity affects both the ability to ensure compliance with the law, to collect information about the population, to extract revenue, and to manage and implement more complex policies.

Model 5 substitutes the electoral democracy measure with a measure of liberal democracy. The liberal democracy index captures aspects beyond the elective, and focuses on among other things the extent to which civil liberties and minority rights are protected as well as whether or not the executive power is constrained by effective checks and balances (Coppedge et al., 2021). I have included this measure in part because I want to make sure that the strong positive relationship between liberal democracies and more capable states depicted in Figure 6.2a does not mean that more capable states are capable *because* they are liberal.

When liberal democracy is introduced, State Capacity loses some of its predictive power. As pointed out by for instance Zakaria (1997) and Fjelde, Knutsen, and Nygård (2021), liberal democracies display both extensive electoral rights, protection of basic liberties, and strong parliaments and judiciaries able to constrain the chief executive. The combination of vertical and horizontal constraints is widely expected to limit the undermining of minority groups' rights, which furthermore should lessen grievances, increase political trust, and facilitate for nonviolent political practices (Fjelde et al., 2021). However, the liberal democracy variable is only statistically significant in predicting conflict onset if State Capacity and natural resource wealth are removed from the model. Although to a somewhat lesser extent, State Capacity still predicts a 36% decrease in the odds of conflict outbreak for every one-unit increase in itself. In terms of predicted probabilities for this model, states that score below -1 on State Capacity have probabilities between 3.2% and 6.2% on average when all other variables are held constant at their means [95% CI

= .003, .129]. In comparison, states that score above 1 have predicted probabilities of conflict outbreak between 1.32% and 1% [95% CI = .00002, .0185].

Once again, these are good news for the new, aggregated State Capacity measure. Some additional diagnostics and robustness checks can be found in Appendix Table D.3, Appendix Table D.4, Appendix Table D.6 and Appendix Table D.5. In D.3, I check whether State Capacity still is able to predict conflict onset when the conflict is either new or if it has only been a year (or more) since the last observation of the conflict (*onset1*). Similarly, D.4 tests whether State Capacity has predictive power when it has been five or more years since the last observation of the conflict (*onset5*). In D.6, I control for ethnic fractionalization. Ethnic fractionalization, obtained from the Historical Index of Ethnic Fractionalization Dataset (Drazanova, 2020), captures longitudinal changes in ethnic diversity by measuring whether individuals in a country tend to belong to the same or a diversity of groups.

Recognizing the important findings of among others Østby (2008), who argue that it is not ethnic heterogeneity *per se* that makes states more prone to conflicts but rather distributional group-asymmetries, highly fractionalized societies are still expected to be more prone to conflict onset (Esteban & Ray, 2008; Drazanova, 2020). For example, Wegenast and Basedau (2014) find that the combination of natural resources and shared identities helps to overcome the collective action problems associated with rebellion, in large due to greed-associated factors such as higher recruitment possibilities, lower opportunity costs and stronger economic incentives. Lastly, as the rate of outbreak was highest in Sub-Saharan Africa, and because most of the top ranking countries on the State Capacity variable are "Western", D.5 controls for regional effects. Two dummies are introduced separately: one for Western Europe and North America, and one for Sub-Saharan Africa.

Capacity performs even better than expected, and yields statistically significant result in every model for every robustness check, except for model 5 in D.3. Even after the Sub-Saharan Africa region is controlled for, together with GDP/capita, population size, number of peace years, cubic splines, resource wealth and electoral democracy, the odds of conflict onset is 41% higher for each one-unit decrease in State Capacity. When ethnic fractionalization is added to this equation, the number

of observations drops to 3.327, number of conflicts to 111 and years covered 1992-2013, but the odds of conflict outbreak is estimated 47% lower for every one-unit increase in State Capacity. When removing the regional dummies from the list of control variables, the predicted probability of conflict onset for countries scoring $\leq$ -1 on State Capacity is 7% when all above-mentioned control variables are held at their means conditional on State Capacity ≤ 1 [95% CI = .047, .093]. In comparison, for countries with State Capacity scores ≥ 1, the predicted probability of conflict onset is 0.5% [95% CI = .001, .009]. All in all, State Capacity remains a strong and consistent predictor. Even in every robustness check, the odds of conflict onset is predicted between 31% and 62% higher for every one-unit decrease in State Capacity.

As a last robustness check, I compared State Capacity's performance in predicting conflict onset with five of the other, closely related measured discussed in the last section, using receiver operating characteristic (ROC) curves. The results can be found in Appendix Figure D.2. In these tests, the sample size was reduced in order to compensate for missingness in some of the variables and achieve comparability across models. Tests were conducted on a limited set of 1,712 observations and 74 conflict onsets for the period 2007-2019, controlling for income per capita and population size. All predictors perform more or less equally well, only with tiny differences in AUC scores. Out of the six predictors, State Capacity was the preferred with an AUC score of .8032, while WGI's Rule of Law performing the "worst" with AUC equal to .7869. However, the differences are minimal. That predicted estimates for State Capacity are similar to other measures' is nevertheless a comforting indication when it comes to State Capacity's validity.

6.4 Summary

In this chapter, I have explored the aggregate measure more thoroughly. After investigating whether "on its face" the measure makes sense and seems like a good translation of the construct, whether it is comparable to other aggregate measures that it theoretically should converge on and how it performs when it is asked to

predict armed conflict onset, I concluded that the measure as presented here has high overall validity and that its power and applicability is significant. It is furthermore suggested that the measure is more comprehensive and that it taps into the core concept of state capacity in a more satisfactory manner than previous attempts. State Capacity is founded on a solid theoretical base, and by drawing on indicators representing all the hypothesized dimensions that are minimally necessary for the modern state to execute its core functions, it is expected to function well as a general-purpose measure with broad utility and relevance.

Thus, the results from this chapter provide additional, validated support for the theoretical argument made that it is more constructive to view the extractive, coercive, and bureaucratic-administrative dimensions of state capacity as mutually interrelated and codependent, not only in theory, but also in practice. The procedures undertaken avoid conflating state capacity with other conceptually distinct aspects to the state, and the measure seems to be a strong predictor of conflict outbreak. This should be investigated more closely with more complex models in further research. The next and last chapter provides a thesis summary and a discussion of results, contributions, limitations and implications.

Chapter 7

Discussion and Conclusion

7.1 Summary and Findings

The main emphasis in this thesis has been to address key conceptual and empirical issues raised by measuring state capacity in studies on intrastate armed conflict. I have argued that current applications of state capacity in quantitative analysis suffer from significant flaws both theoretically and empirically. This is first of all because prevailing proxies often are poorly conceptualized and not seldom employed without taking construct validity into account. Second of all, current indicators tend to drift from a core theoretical focus on the material and organizational aspects to state that are key for its abilities to implement desired policies and thus avoid conflict outbreak. Instead, the concept is conflated with other, conceptually distinct state features like economic development, regime type and redistribution efforts, or the political decision to prioritize suchlike normative outcomes.

Third, researchers regularly describe state capacity as involving only the state functions that are relevant to their particular inquiry. As a result, conflicting conceptualizations have led a great deal of state capacity proxies to either be aggregated into questionable catch-all indices, or to be broken down to highly diaggregated sets of state functions, leaving a confusing array of understandings of statehood and insufficient attention to how different dimensions are interrelated and codependent. This thesis has dealt with this methodological inconsistency in three ways.

In chapter 3, I assessed competing conceptualizations and measurements of state capacity on the basis of construct validity, focusing criticism on the threats to inference that arise when a measure either does a poor job capturing the theoretical construct, or fails to discern between competing causal mechanisms. Subsequently, I scrutinized in detail the three dimensions expected to be mutually interlinked and the indicators by which they are operationalized. These 'general underpinnings' of state capacity relate to 1) the state's coercive capacity to preserve its borders, to contain threats, and to maintain internal order, 2) the bureaucratic-administrative sophistication necessary to collect, monitor and implement data, information and policy, to reach citizens, and to conduct decision-making in an impartial manner, and 3) the state's ability to extract revenue in order to support all other activities.

Taking as point of departure this reconceptualization, chapter 5 explored the interrelationship between 20 carefully selected indicators using principal factor analysis. The results provide strong support for the proposition that the dimensions are codependent and hard to distinguish empirically, and low underlying dimensionality led one factor to explain close to 67% of the variance in the measures. To a large extent, this factor, termed State Capacity, captures the bundle of qualities that were hypothesized to make up the state's ability to carry out its core functions: it favors bureaucratic, technologically advanced countries capable of extracting information and taxes that require higher levels of societal penetration, and high-ranking states according to this factor are generally able to ensure compliance with the law and internal order by other means than coercive.

The only factor to represent a distinct state capacity dimension was factor 3, on which the indicators representing devotion to the military and natural resource dependence load heavily. This was interpreted as a confirming sign that it is not the number of dollars to GDP put into defence nor the size of the security force that matter for the level of coercive capacity, but rather administrative organization and technological sophistication (Hendrix, 2010; Hanson & Sigman, 2020). Instead, it is likely that larger proportions of military personnel and expenditure reflect symptoms of insurgent mobilization and increased threat perceptions (and thus compensation through investments in repression) or escalation of current conflicts. The fact that

Libya, Syria, Yemen and Somalia ranked among the top-15 countries in 2019 support the argument that state presence and military devotion often *follows* and not *precedes* violent conflict, rejecting the notion that larger armies represent capacity of deterrence and hindering rebellion (Henderson & Singer, 2000; Gupta et al., 2001; Koren & Sarbahi, 2017).

That most of the high-ranking countries in factor 3 also are oil-rich, resource dependent and highly autocratic support the theory and empirics holding that 'petroleum generates violence' (de Soysa, 2000; Reno, 2000; P. Collier & Hoeffler, 2004; Bates, 2008; Besley & Persson, 2010; Ross, 2015; de Soysa et al., 2021). Regardless of which causal mechanism connecting the resource-conflict correlation one might prefer, a combined logic is plausible. By providing lootable income to private actors, bolstering "shadow states", increasing corruption, decreasing incentives to invest in state building, lowering opportunity costs of capture and rebellion, motivating ethnic prospects of independence, and encouraging greed-driven conflict, higher levels of natural resource dependence can comfortably be interpreted as being associated with lower state capacity.

Table 7.1 illustrates the results from a simple logistic regression in which the estimated factor 3 predicts conflict onset, controlling for income, population size, State Capacity and peace years. As theory would predict, the combination of natural resource dependence and more devotion to the military is strongly associated with an increased risk of conflict onset. In fact, for every one-unit increase in factor 3, the estimated odds of conflict outbreak are multiplied by 1.32. Although this finding is reasonable, one should expect a certain (large) degree of endogeneity also in this relationship. Again, the association may very well be driven by the fact that highly militarized states already are mired in conflict, causing the estimate to be biased upwards. This uncertainty is also evident in the natural resource abundance logic. Instead of measuring natural resource dependence *per se*, resource intensive production could simply be the result of poor economic performance and lack of industrialization for reasons other than resource abundance. This makes it hard to interpret relationships between resource abundance and frequency of conflicts, as both might be boosted by the state's poverty or lack of development (Ross, 2015).

TABLE 7.1: The effect of factor 3 on conflict onset, 1992-2017

	Odds Ratio	*Robust Std.Err.*	*z*	*p*	*[95% Conf.*	*Interval]*
Factor 3[a]	1.322026	.1437801	2.57	0.010	1.068232	1.636117
(log) GDP/capita[a]	.5600785	.2049842	-1.58	0.113	.2733481	1.147577
(log) Population[a]	3.379833	.68091	6.04	0.000	2.277246	5.016265
State Capacity[a]	.6002787	.1354285	-2.26	0.024	.385757	.9340971
Peace Years	.9637571	.0141431	-2.52	0.012	.936432	.9918795
Constant	.0756721	.0980143	-1.99	0.046	.0059761	.9581965
N						4227

Notes: (a) lagged one year; standard errors are robust, clustered by country-ID

Moving on, chapter 6 deep-dived into the new State Capacity measure and investigated whether it behaves in the expected manner. The measure was examined in terms of its face validity, content validity, convergent validity and predictive validity. Its face validity was inspected by assessing whether or not the measure on the surface "makes sense". In sum, I have argued that the measure should be satisfactory in terms of face- and content validity unless important measures are overlooked. The highest-scoring countries correspond to the countries one would expect to rank high, the measure is successful in distinguishing itself from measures related to regime type and economic development, and its temporal variation is consistent with expected trends in that most countries' state capacities increase over time. Furthermore, naturally, it correlates highly with its base indicators and draws from all three theorized dimensions, confirming that the measure should work well as a general-purpose measure of relevance far beyond this particular inquiry.

In terms of convergent validity, pairwise correlations between State Capacity and other measures that it theoretically should converge on confirm that the measure functions in the expected manner. It is for example highly correlated with the Worldwide Governance Indicators' measures of Government Effectiveness $r = .93$, Rule of Law $r = .91$ and Regulatory Quality $r = .89$, Fund for Peace's Fragile States Index $r = -.87$, and Center for Systemic Peace's State Fragility Index $r = -.79$. As was illustrated in Figure 6.4, State Capacity also behaves in the expected manner when compared to measures it is harder to distinguish it from. Nevertheless,

it is a good sign that the comparisons also yield differences in the measures, as they indeed measure different things. For example, the Fragile States Index mix measures of state legitimacy, security apparatus, provision of public services and group grievances in their aggregation procedure. On the other hand, while WGI's Government Effectiveness draws upon a multitude of different indicators like quality of railways, public school coverage and welfare, it also includes some of the same indicators as State Capacity, such as ICRG's measure of bureaucratic quality[13]. This might to some extent explain why they are so highly correlated.

When it comes to interchangeability and predictive ability, chapter 6 demonstrated the utility of the new aggregate measure using the level of State Capacity as a predictor for intrastate armed conflict onset. The results from the replication of Fearon & Laitin's (2003) first two models dispute their argument that per capita income is a proxy for a state's overall financial, administrative, police, and military capabilities [p.80]. When State Capacity is added to the models, GDP/capita drops out of significance. Furthermore, in both models, State Capacity was a significantly strong predictor, and estimated that for each one-unit decrease in State Capacity, the odds of observing civil war outbreak between 1992 and 1999 increased by 84.6%. However, as analysis only is conducted for eight years, the results are not necessarily generalizable for a longer period of time and should therefore be interpreted with caution. Nevertheless, the findings reveal that income per capita is a flexible measure, inadequate in capturing the broad range of state capacities necessary to prevent conflict.

Lastly, chapter 6 further tested the utility of State Capacity by running a logistic regression, testing whether State Capacity significantly can predict conflict onset in the time period of 1992-2017 even after controlling for other, important aspects to the state. In each of the five models, State Capacity was a statistically significant predictor, estimating that each one-unit decrease in State Capacity was associated with a 36%(model 5) to 57%(model 1) increase in the odds of conflict outbreak. Across the models, the predicted probabilities that a state scoring -2 on State Capacity would observe intrastate conflict onset when all other independent

[13]See `https://info.worldbank.org/governance/wgi/Home/Documents`

variables were held at their means ranged from 10% (model 1) to 5% (model 5). In comparison, this estimate was between 0.42% (model 1) and 0.84% (model 5) for states scoring 2 on State Capacity.

7.2 Contributions

In conclusion, this thesis contributes to the literature in a number of ways. First of all, revisiting the state of state capacity's conceptualizations, operationalizations and construct validity is an important task per se. As repeatedly touched upon, research can only be relevant, reliable and systematically unbiased if data and approaches to modeling are enhanced, and this thesis has contributed with a critical validity assessment of current state capacity proxies commonly employed in research on armed conflict. It has furthermore attempted to reconceptualize the concept in a way that avoids conflating it with other conceptually distinct concepts, only capturing the very core state functions that are minimally necessary for the modern state to implement desired (conflict-reducing) policies. This reconceptualization provides the basis for a measurement strategy that can be applied not only to studies on peace and conflict, but to all comparative cross-national research concerned with the causes and consequences of state capacity, which may include state capacity's effect on public goods provision, health- and schooling outcomes, drug control and crime rates, and (illegal) economic activity.

Thus, the thesis represent one of the first steps towards bridging an important methodological gap. Second, the thesis constructs a new measure with high demonstrated validity and utility as a strong and consistent predictor of intrastate conflict outbreak. The fact that the measure predicts similar results as other aggregate measures attempting to capture the same concept, and that it at the same time is able to distinguish itself from conceptually distinct concepts, suggests that the measure is indeed capturing something that these other measures are not. Founded on a solid theoretical base in which construct validity is critically assessed, the measure successfully draws upon all three theorized dimensions and should capture the underlying concept more comprehensively than previous attempts that are either

focused on a single disaggregated indicator or dimension, or aggregated into catch-all indices. In sum, this thesis has contributed to the field with a potentially more valid, theoretically consistent and reliable measurement of state capacity, though bearing in mind the difficulties and imperfectness of statistically measuring a social construct.

7.3 Limitations and Implications

Despite important contributions, there are some limitations to this thesis worth elaborating on. The fist concerns the conceptual fitting of and potential biases in the indicators that make up the aggregate State Capacity measure. As touched upon in chapter 4, the choice of variables to include in EFA and a theoretically sound base matter substantially for the identification of 'true' common factors. The variables that ended up being included are, to the best of my knowledge, the most conceptually fitting, frequently employed, and recognized measures relating to the three identified dimensions of state capacity. Multiple other measures were assessed but excluded, either because of conceptual inconsistency, lacking representativeness, or because the causal mechanism linking the measure to state capacity could be highly disputed. I have worked to overcome the omitted variable bias by first setting clear theoretical boundaries in chapter 3, thereafter by assessing a broad variety of measures, and subsequently by being transparent about selection criteria. Although unlikely, there is a possibility that important indicators that are both theoretically and empirically important for either of state capacity's three dimensions have been overlooked, or that the included measures could have been substituted by proxies that are measured in a more fitting way or that are from even more reliable sources. Nevertheless, the obtained factor works in a satisfactory manner, suggesting that the included variables are adequate.

Another concern in this regard is endogeneity between the different indicators and the outcome they are trying to predict in chapter 6. Although conflicts arise from a number of complex factors that often work in tandem, one should accept the finding that more capable states should be better able to avoid outbreaks. Nev-

ertheless, the interplay between conflict and state capacity is hard to unpack, and there is currently no available way to guarantee that included measures are entirely independent. They might capture aspects of each other on the one hand and aspects directly related to conflict on the other. In other words, it is possible that the state capacity measure to some extent reflect peaceful vs. conflict-affected countries rather than strong and weak state capacity per se. This is especially true for the included expert-coded assessments like ICRG's measure of law and order, but also for the measures capturing military devotion and natural resource dependence. Expert assessments should be highly sensitive to information about increased political stability, growing unrest and historical knowledge. Thus, one can never be completely certain that these measures are unaffected by subjective assessments of exogenous factors to the variable they are measuring. However, because the indicator is part of an aggregate, weighted measure, the problem is likely to be smaller than would be the case if this variable was used as a single predictor alone. I have at least partly addressed this problem by carefully evaluating the measures beforehand and lagging State Capacity in the regressions. The endogeneity problem of the measures capturing military devotion and natural resource dependence should not be of substantial effect in this study, as the variables are only weakly correlated with the main independent variable of interest and does not load significantly on its factor.

However, the perhaps biggest limitation to the thesis relates to missing data and multiple imputation. I have outlined this issue thoroughly in chapter 4, but it deserves to be stressed again that it is far from ideal to measure state capacity - which is already a highly uncertain construct - based on data in which some variables contain a relatively high proportion of missingness. Dealing with missingness in factor analyses remains an unsolved challenge, and has caused for instance Hendrix' (2010) analysis to be relatively restricted, with only 1,610 country-years observations for 101 countries from 1984-1999. In my case, complete case analysis would lead to an extreme loss of information and severely biased estimates. Therefore, in order to compensate for missingness the best way possible, predictive mean matching by multiple imputation using R's MICE package was employed.

Furthermore, the fact that I averaged the ten imputation models can be sub-

ject of criticism. When averaging imputations prior to running the factor analysis, information about uncertainty around estimates is lost, leading the imputed values to be interpreted as real and the correlations to be somewhat inflated. There are numerous ways in which this could be done differently, but since there is no general consensus on which approach is the best when used in EFA, I nevertheless proceeded with the averaging method. Although not a perfect solution, the kernel density plots yield results that the averaged imputed estimates are adequate and satisfactory for this particular study: the imputed values fit accurately within the natural range of the observed values, and the imputed distributions generally match the observed well. Moreover, the combination of EFA and multiple imputation should not be avoided simply because it is demanding and unsolved, rather, one should encourage future research to pay more attention to these particular challenges. In any way, the readers of this thesis should be aware of the limitations - especially those related to endogeneity in measures and factors obtained using imputed data - and interpret the results accordingly.

7.4 Moving Further

Ultimately, this thesis represents one of the first contributions towards bridging an important methodological gap in the state capacity-armed conflict literature and provides the basis for a measurement strategy that can be applied not only to studies on peace and conflict, but to all comparative cross-national research concerned with the causes and consequences of state capacity. With a growing interest in the definitions and measurements of state capacity (recently addressed by e.g. Hendrix, 2010; Lindvall & Teorell, 2016; Centeno et al., 2017; Berwick & Christia, 2018; Hanson & Sigman, 2020), approaches to modeling are likely to be intensely improved over the next years. The main takeaway from this thesis is that more attention should be paid to construct validity when choosing between measures to include in quantitative analysis, and that such measures should, in the best way possible, adjudicate between competing causal mechanisms. The approach undertaken here is simply a beginning. Yet, an important one.

References

Abayomi, K., Gelman, A., & Levy, M. (2008). Diagnostics for Multivariate Imputations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *57*(3), 273-291. doi: 10.1111/j.1467-9876.2007.00613.x

Adcock, R., & Collier, D. (2001). Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *The American Political Science Review*, *95*(3), 529–546. doi: 10.1017/S0003055401003100

Anderson, B. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.

Anderson, C., & Carter, J. (2019). *Principles of Conflict Economics: The Political Economy of War, Terrorism, Genocide, and Peace*. Cambridge: Cambridge University Press.

Arreguín-Toft, I. (2005). *How the Weak Win Wars: A Theory of Asymmetric Conflict*. Cambridge: Cambridge University Press.

Azur, M., Stuart, E., Frangakis, C., & Leaf, P. (2011). Multiple Imputation by Chained Equations: What is it and how does it work? *International journal of methods in psychiatric research*, *20*(1), 40-49. doi: 10.1002/mpr.329

Bakaki, Z. (2020). The Joint Effect of International and Domestic-Level State Capacity on Civil War Risk. *Social Sciences*, *9*(5). doi: 10.3390/socsci9050066

Bates, R. (2008). *When Things Fell Apart: State failure in late-century Africa*. Cambridge: Cambridge University Press.

Bates, R. (2009). *Prosperity and Violence* (2nd ed.). W.W. Norton Company.

Bates, R., Greif, A., & Singh, S. (2002). Organizing violence. *Journal of Conflict Resolution*, *46*(5), 599-628. doi: 10.1177/002200202236166

Beck, N., Katz, J., & Tucker, R. (1998). Taking Time Seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science*, *42*(4), 1260-1288. doi: 10.2307/2991857

Bemelmans-Videc, M.-L., Rist, R., & Vedung, E. (1998). *Carrots, Sticks Sermons: Policy Instruments and their Evaluation.* New Brunswick, N.J., US: Transaction Publishers.

Bennett, A. (2004). Case Study Methods: Design, Use, and Comparative Advantages. In D. Sprinz & Y. Wolinsky-Nahmias (Eds.), *Models, numbers cases: Methods for studying international relations* (p. 19-55). United States of America: The Univeristy of Michigan Press.

Berwick, E., & Christia, F. (2018). State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate. *Annual Review of Political Science*, *21*, 71-91. doi: 10.1146/annurev-polisci-072215-012907

Besley, T., & Persson, T. (2009). The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review*, *99*(4), 1218–1244. doi: 10.1257/aer.99.4.1218

Besley, T., & Persson, T. (2010). State Capacity, Conflict, and Development. *Econometrica*, *78*(1), 1-34. doi: 10.2307/25621395

Blalock, H. (1979). The Presidential Address: Measurement and Conceptualization Problems: The Major Obstacle to Integrating Theory and Research. *American Sociological Review*, *44*(6), 881-894. doi: 10.2307/2094714

Bockstette, V., Chanda, A., & Putterman, L. (2002). States and Markets: The Advantage of an Early Start. *Journal of Economic growth*, *7*(4), 347-369. doi: 10.1023/A:1020827801137

Braithwaite, A. (2010). Resisting Infection: How State Capacity Conditions Conflict Contagion. *Journal of Peace Research*, *47*(3), 311–319. doi: 10.1177/0022343310362164

Brambor, T., Goenaga, A., Lindvall, J., & Teorell, J. (2020). The Lay of the Land: Information Capacity and the Modern State. *Comparative Political Studies*, *53*(2), 175-213. doi: 10.1177/0010414019843432

Buhaug, H. (2010). Dude, Where's My Conflict?: LSG, Relative Strength, and

the Location of Civil War. *Conflict Management and Peace Science*, *27*(2), 107–128. doi: 10.1177/0738894209343974

Buhaug, H., Gates, S., & Lujala, P. (2009). Geography, Rebel Capability, and the Duration of Civil Conflict. *Journal of Conflict Resolution*, *53*(4), 544-569. doi: 10.1177/0022002709336457

Buhaug, H., & Gleditsch, K. (2008). Contagion or Confusion? Why Conflicts Cluster in Space. *International Studies Quarterly*, *52*(2), 215–233. doi: 10.1111/j.1468-2478.2008.00499.x

Bustikova, L., & Corduneanu-Huci, C. (2017). Patronage, Trust, and State Capacity: The Historical Trajectories of Clientelism. *World Politics*, *69*(2), 277-326. doi: 10.1017/S0043887116000265

Börzel, T., & Risse, T. (2016). Dysfunctional State Institutions, Trust, and Governance in Areas of Limited Statehood. *Regulation & Governance*, *10*, 149-160. doi: 10.1111/rego.12100

Carment, D., Landry, J., & Samy, Y. (2014). State failure, development, and international security. In E. Newman & K. DeRouen (Eds.), *Routledge handbook of civil wars* (p. 334–345). New York: Routledge.

Carmines, E., & Zeller, R. (1979). *Reliability and validity assessment: Quantitative applications in the social sciences*. Thousand Oaks, CA: Sage. doi: 10.4135/9781412985642

Cederman, L.-E., Gleditsch, K., & Buhaug, H. (2013). *Inequality, grievances, and civil war*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139084161

Cederman, L.-E., & Vogt, M. (2017). Dynamics and Logics of Civil War. *Journal of Conflict Resolution*, *61*(9), 1992-2016. doi: 10.1177/0022002717721385

Centeno, M., Kohli, A., Yashar, D., & Mistree, D. (2017). *States In the Developing World*. Cambridge: Cambridge University Press.

Checkel, J. (2008). Bridging the Gap? Connecting Qualitative and Quantitative Methods in the Study of Civil War. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research*, *6*(1), 13–14.

Child, D. (1990). *The Essentials of Factor Analysis*. London & New York: Continuum International Publishing Group.

Cingolani, L. (2013, May). *The State of State Capacity: a review of concepts, evidence and measures* (UNU-MERIT Working Paper Series). Maastricht Graduate School of Governance: Maastricht Economic and Social Research Institute on Innovation and Technology.

Clayton, G. (2014). Quantitative and econometric methodologies. In E. Newman & K. DeRouen (Eds.), *Routledge handbook of civil wars* (p. 28–40). New York: Routledge.

Collier, D., & Mahoney, J. (1996). Insights and Pitfalls: Selection Bias in Qualitative Research. *World Politics*, *49*(1), 56–91. doi: 10.1353/wp.1996.0023

Collier, P., Elliott, V., Hegre, H., Hoeffler, A., Reynal-Querol, M., & Sambanis, N. (2003). *Breaking the conflict trap: Civil war and development policy*. Oxford: Oxford University Press. Retrieved from `http://econ.worldbank.org/prr/CivilWarPRR/`

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, *56*(4), 563–595. doi: 10.1093/oep/gpf064

Conrad, J., Kim, H.-C., & Souva, M. (2013). Narrow interests and military resource allocation in autocratic regimes. *Journal of Peace Research*, *50*(6), 737-750. doi: 10.1177/0022343313498885

Coppedge, M., Gerring, J., Knutsen, C., Lindberg, S., Teorell, J., Altman, D., ... Ziblatt, D. (2021). *V-Dem Codebook v11.1.* Varieties of Democracy (V-Dem) Project. Retrieved from `https://www.v-dem.net/en/data/data/v-dem-dataset-v111/`

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. doi: 10.1037/h0040957

Cunningham, D., Gleditsch, K., & Salehyan, I. (2009). It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome. *Journal of Conflict Resolution*, *53*(4), 570–597. doi: 10.1177/0022002709336458

Cunningham, D.E. and Lemke, D. (2014). Beyond civil war: A quantitative examination of causes of violence within countries. *Civil Wars*, *16*(3), 328–345. doi:

10.1080/13698249.2014.976356

Dahl, R. (1957). The Concept of Power. *Behavioral Science*, *2*(3), 201–215. doi: 10.1002/bs.3830020303

Dal Bó, E., & Dal Bó, P. (2011). Workers, Warriors, and Criminals: Social Conflict in General Equilibrium. *Journal of the European Economic Association*, *9*(4), 646-677. doi: 10.1111/j.1542-4774.2011.01025.x

Davies, J. (1962). Toward a theory of revolution. *American sociological review*, *27*(1), 5-19. doi: https://doi.org/10.2307/2089714

DeRouen, K., Ferguson, M., Norton, S., Park, Y., Lea, J., & Streat-Bartlett, A. (2010). Civil war peace agreement implementation and state capacity. *Journal of Peace Research*, *47*(3), 333–346. doi: 10.1177/0022343310362169

DeRouen, K., & Sobek, D. (2004). The Dynamics of Civil War Duration and Outcome. *Journal of Peace Research*, *41*(3), 303–320. doi: 10.1177/0022343304043771

DeRouen, K., & Sobek, D. (2016). State Capacity, Regime Type, and Civil War. In T. Mason & S. Mitchell (Eds.), *What do we know about civil wars?* (p. 59-74). United Kingdom: Rowman and Littlefield.

de Soysa, I. (2000). The Resource Curse: Are Civil Wars Driven by Rapacity or Paucity? In M. Berdal & D. Malone (Eds.), *Greed & grievance: Economic agendas in civil wars* (p. 113-136).

de Soysa, I., Vadlamannati, K., & Finseraas, H. (2021). Group Grievances & Civil War: Some Theory and Empirics on Competing Mechanisms, 1990-2017. , Presented at ISA 2021 (Forthcoming).

de Winter, J. C., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*(4), 695-710. doi: 10.1080/02664763.2011.610445

Di John, J. (2010). The Concept, Causes and Consequences of Failed States: A Critical Review of the Literature and Agenda for Research with Specific Reference to Sub-Saharan Africa. *The European Journal of Development Research*, *22*(1), 10-30. doi: 10.1057/ejdr.2009.44

Dray, S., & Josse, J. (2015). Principal Component Analysis with Missing Values: a Comparative Survey of Methods. *Plant Ecology*, *216*(5), 657-667. doi: 10.1007/s11258-014-0406-z

Drazanova, L. (2020). Introducing the Historical Index of Ethnic Fractionalization (HIEF) Dataset: Accounting for Longitudinal Changes in Ethnic Diversity. *Journal of Open Humanities Data*, *6*(1). doi: 10.5334/johd.16

Esteban, J., & Ray, D. (2008). Polarization, Fractionalization and Conflict. *Journal of peace Research*, *45*(2), 163-182. doi: 10.1177/0022343307087175

Evans, p., & Rauch, J. (1999). Bureaucracy and Growth: A Cross-National Analysis of the Effects of "Weberian" State Structures on Economic Growth. *American sociological review*, 748-765. doi: 10.2307/2657374

Fabrigar, L., Wegener, D., MacCallum, R., & Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272-299. doi: 10.1037/1082-989X.4.3.272

Fearon, J. (1995). Rationalist Explanations for War. *International Organization*, *49*(3), 379–414. doi: 10.1017/S0020818300033324

Fearon, J., & Laitin, D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, *97*(1), 75–90. doi: 10.1017/S0003055403000534

Fisunoglu, A., Kang, K., Arbetman-Rabinowitz, M., & Kugler, J. (2011). *Relative Political Capacity Dataset (Version 2.4) August 2020.* Harvard Dataverse. doi: 10.7910/DVN/NRR7MB

Fjelde, H., & de Soysa, I. (2009). Coercion, Co-optation, or Cooperation? State Capacity and the Risk of Civil War, 1961–2004. *Conflict Management and Peace Science*, *26*(1), 5–25. doi: 10.1177/0738894208097664

Fjelde, H., Knutsen, C., & Nygård, H. (2021). Which Institutions Matter? Re-Considering the Democratic Civil Peace. *International Studies Quarterly*, *65*(1), 223-237. doi: 10.1093/isq/sqaa076

Fukuyama, F. (2011). *The origins of political order: From prehuman times to the French Revolution.* New York: Farrar, Straus and Giroux.

Fukuyama, F. (2013). What is Governance? *Governance*, *26*(3), 347-368. doi: 10.1111/gove.12035

Fukuyama, F. (2014). *Political Order and Political Decay: From the Industrial Revolution to the Globalization of Democracy.* London: Profile Books.

Gates, S., Nygård, H., & Strand, H. (2012). Development Consequences of Armed Conflict. *World Development*, *40*(9), 1713–1722. doi: 10.1016/j.worlddev.2012.04.031

Gates, S., Nygård, H., & Trappeniers, E. (2016). Conflict Recurrence. *PRIO Conflict Trends*, *2*. Retrieved from `https://www.prio.org/Publications/Publication/?x=9056`

Geddes, B. (1996). *Politician's Dilemma: Building State Capacity in Latin America.* Berkley, L.A. and London: University of California Press.

Giraudy, A., Luna, J., & Mistree, D. (2017). Unpacking the State's Uneven Territorial Reach: Evidence from Latin America. In M. Centeno, A. Kohli, & D. Yashar (Eds.), *States in the developing world* (p. 93–120). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781316665657.005

Gjerløw, H., Knutsen, C., Wig, T., & Wilson, M. (2021). *One Road to Riches? How state building and democratization affect economic development.* (Unpublished)

Gleditsch, K., Dorussen, H., Metternich, N., & Ruggeri, A. (2012). *Transnational Dimensions and the Myth of Civil Wars as National Events.* Retrieved from `https://www.prio.org/Publications/Publication/?x=7261`

Gleditsch, K., & Ruggeri, A. (2010). Political opportunity structures, democracy, and civil war. *Journal of Peace Research*, *47*(3), 299–310. doi: 10.1177/0022343310362293

Gleditsch, K., & Ward, M. (1999). A revised list of independent states since the congress of Vienna. *International Interactions*, *25*(4). doi: 10.1080/03050629908434958

Gleditsch, N.-P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research*, *39*(5). doi: 10.1177/0022343302039005007

Gomide, A., Pereira, A., & Machado, R. (2018, June). The Concept of State Capacity and its Operationalization in Empirical Research. In *International*

workshops on public policy, international public policy association (p. 2-20). University of Pittsburgh.

Goodwin, J., & Skocpol, T. (1989). Explaining Revolutions in the Contemporary Third World. *Politics & Society*, *17*(4), 489–509. doi: 10.1177/003232928901700403

Grävingholt, J., Ziaja, S., & Kreibaum, M. (2015). Disaggregating State Fragility: a Method to Establish a Multidimensional Empirical Typology. In R. Gisselquist (Ed.), *Fragility, Aid, and State-building: Understanding Diverse Trajectories* (p. 13-30). London and New York: Routledge.

Gupta, S., de Mello, L., & Sharan, R. (2001). Corruption and Military Spending. *European Journal of Political Economy*, *17*(4), 749-777. doi: 10.1016/S0176-2680(01)00054-4

Gurr, T. (1970). *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Hanson, J., & Sigman, R. (2020). Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research.. (Unpublished) doi: 10.1086/715066

Harbers, I. (2015). Taxation and the Unequal Reach of the State: Mapping State Capacity in Ecuador. *Governance*, *28*(3), 373-391. doi: 10.1111/gove.12117

Harris, R. (1970). *Death of a Revolutionary: Che Guevara's Last Mission*. New York: Norton.

Hegre, H. (2018). Civil Conflict and Development. In N. van de Walle & C. Lancaster (Eds.), *Oxford university press handbook on the politics of development*. Oxford: Oxford University Press.

Hegre, H., Ellingsen, T., Gates, S., & Gleditsch, N.-P. (2001). Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816-1992. *The American Political Science Review*, *95*(1), 33–48. doi: 10.1017/S0003055401000119

Hegre, H., & Nygård, H. (2014). Governance and conflict relapse. *Journal of Peace Research*, *47*(3), 299–310. doi: 10.1177/0022002713520591

Hegre, H., Strand, H., Gates, S., & Nygård, H. (2011). The conflict trap. In *Apsa 2011 annual meeting paper* (p. 1–31). Retrieved from `https://ssrn.com/`

abstract=1900679

Henderson, E., & Singer, D. (2000). Civil War in the Post-Colonial World, 1946-92. *Journal of Peace Research*, *37*(3), 275-299. doi: 10.1177/0022343300037003001

Hendrix, C. (2010). Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict. *Journal of Peace Research*, *47*(3), 273–285. doi: 10.1177/0022343310361838

Hendrix, C. (2011). Head for the Hills? Rough Terrain, State Capacity, and Civil War Onset. *Civil Wars*, *13*(4), 345–370. doi: 10.1080/13698249.2011.629863

Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, NJ: Princeton University Press.

Herz, J. (1951). *Political Realism and Political Idealism*. Chicago: University of Chicago Press.

Hillesund, S., Bahgat, K., Barrett, G., Dupuy, K., Gates, S., Nygård, H., ... Østby, G. (2018). Horizontal Inequality and Armed Conflict: A Comprehensive Literature Review. *Canadian Journal of Development Studies*, *39*(4), 463-480. doi: 10.1080/02255189.2018.1517641

Hobbes, T. (2008). *Leviathan* (J. Gaskin, Ed.). New York: Oxford University Press.

Hobsbawm, E. J. (2012). *Nations and Nationalism Since 1780: Programme, Myth, Reality*. New York: Cambridge university press.

Honaker, J., & King, G. (2010). What to Do about Missing Values in Time-Series Cross-Section Data. *American Journal of Political Science*, *54*(2), 561-581. doi: 10.1111/j.1540-5907.2010.00447.x

Horowitz, D. (1985). *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.

Howell, L. (2018). ICRG Methodology. Retrieved from https://www.prsgroup.com/wp-content/uploads/2018/01/icrgmethodology.pdf

Huntington, S. (1968). *Political Order in Changing Societies*. New Haven and London: Yale University Press. doi: 10.4324/9780203788332

Högbladh, S. (2020). *UCDP GED Codebook version 20.1.* Department of Peace and Conflict Research: Uppsala University. Retrieved from https://ucdp.uu.se/

downloads/

Jakobsen, T., de Soysa, I., & Jakobsen, J. (2013). Why do poor countries suffer costly conflict? Unpacking per capita income and the onset of civil war. *Conflict Management and Peace Science*, *30*(2), 140-160. doi: 10.1177/0738894212473923

Jervis, R. (1976). *Perception and Misperception in International Politics: New Edition*. Princeton University Press.

Jervis, R. (1978). Cooperation Under the Security Dilemma. *World Politics*, *30*(2), 167-214. doi: 10.2307/2009958

Johnson, D., & Young, R. (2011). Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations. *Journal of Marriage and Family*, *73*(5), 926-945. doi: 10.1111/j.1741-3737.2011.00861.x

Kaiser, H. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and psychological measurement*, *20*(1), 141-151. doi: 10.1177/001316446002000116

Kaplan, S. (2008). *Fixing Fragile States: A New Paradigm for Development: A New Paradigm for Development*. Westport: Praeger Security International.

Kaufmann, C. (1996). Possible and Impossible Solutions to Ethnic Civil Wars.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2003). *Governance matters III: Governance Indicators for 1996-2002* (Working Paper No. 3106). World Bank.

Keen, D. (2008a). *Complex emergencies*. Cambridge: Polity Press.

Keen, D. (2008b). Complex Emergencies: David Keen Responds. Retrieved from https://africanarguments.org/2008/06/complex-emergencies-david-keen-responds/

King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review*, *95*(1), 49-69. doi: 10.1017/S0003055401000235

King, G., Keohane, R., & Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. United Kingdom: Princeton University Press.

Kirk, J., & Miller, M. (1986). *Reliability and Validity in Qualitative Research*.

London: Sage.

Koren, O., & Sarbahi, A. (2017). State Capacity, Insurgency, and Civil War: A Disaggregated Analysis. *International Studies Quarterly*, *62*, 274–288. doi: 10.1093/isq/sqx076

Kugler, J. (2018). *Political Capacity and Economic Behavior*. New York: Routledge Press.

Kugler, J., & Lemke, D. (1996). *Parity and War: Evaluations and Extensions of the War Ledger*. University of Michigan Press.

Levi, M. (1988). *Of Rule and Revenue*. Berkley, CA: University of California Press.

Levitsky, S., & Way, L. (2012). Beyond Patronage: Violent Struggle, Ruling Party Cohesion, and Authoritarian Durability. *Perspectives on Politics*, *10*(4), 869-889. doi: 10.1017/S1537592712002861

Levy, J. (1987). Declining Power and the Preventive Motivation for War. *World Politics*, *40*(1), 82–107. doi: 10.2307/2010195

Lieberman, E. (2002). Taxation Data as Indicators of State-Society Relations: Possibilities and Pitfalls in Cross-National Research. *Studies in Comparative International Development*, *36*(4), 89-115. doi: 10.1007/BF02686334

Lindvall, J., & Teorell, J. (2016, May). *State Capacity as Power: A Conceptual Framework* (STANCE Working Paper Series No. 1). Lund University: Department of Political Science.

Linz, J., & Stepan, A. (1996). *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and post-communist Europe*. Baltimore: John Hopkins University Press.

Little, R. (1988). Missing-Data Adjustments in Large Surveys. *Journal of Business & Economic Statistics*, *6*(3), 287-296. doi: 10.2307/1391878

Mahoney, J., & Goertz, G. (2006). A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research. *Political Analysis*, *14*(3), 227–249. doi: 10.1093/pan/mpj017

Mann, M. (1984). The Autonomous Power of the State: Its Origins, Mechanisms and Results. *European Journal of Sociology/Archives Européennes de Sociologie/Europäisches Archiv für Soziologie*, *25*(2), 185-213. doi: 10.1017/

S0003975600004239

Mann, M. (1993). *The Sources of Social Power*. Cambridge: Cambridge University Press.

Marquardt, K. (2020). How and How Much Does Expert Error Matter? Implications for Quantitative Peace Research. *Journal of Peace Research*, *57*(6), 692-700. doi: 10.1177/0022343320959121

Mason, T., & Greig, M. (2017). State Capacity, Regime Type, and Sustaining the Peace after Civil War. *International interactions*, *43*(6), 967-993. doi: 10.1080/03050629.2017.1260012

Mason, T., Weingarten, J., & Fett, P. (1999). Win, Lose, or Draw: Predicting the Outcome of Civil Wars. *Political Research Quarterly*, *52*(2), 239-268. doi: 10.1177/106591299905200201

Michael, B. (2014). *Statemaking and Territory in South Asia: Lessons from the Anglo-Gorkha War, 1814-1816*. London, New York & Delhi: Anthem Press.

Migdahl, J. (1988). *Strong Societies and Weak States*. Princeton University Press.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis* (An Expanded Sourcebook (2nd ed.). United Kingdom: SAGE Publications, Inc.

Mishra, S. (2016). The Colonial Origins of Territorial Disputes in South Asia. *The Journal of Territorial and Maritime Studies*, *3*(1), 5-23. Retrieved from `https://www.jstor.org/stable/26664123?seq=1#metadata_info_tab_contents`

Moravcsik, A. (2014). Trust, but Verify: The Transparency Revolution and Qualitative International Relations. *Security Studies*, *23*(4), 663–688. doi: 10.1080/09636412.2014.970846

Morris, T., White, I., & Royston, P. (2014). Tuning multiple imputation by predictive mean matching and local residual draws. *BMC medical research methodology*, *14*(1), 1-13. doi: 10.1186/1471-2288-14-75

Muller, E., & Weede, E. (1990). Cross-National Variation in Political Violence: A Rational Action Approach. *The Journal of Conflict Resolution*, *34*(4), 624-651. doi: 10.2307/174182

Nassiri, V., Lovik, A., Molenberghs, G., & Verbeke, G. (2018). On using multi-

ple imputation for exploratory factor analysis of incomplete data. *Behavior Research Methods*, *50*, 501-517. doi: 10.3758/s13428-017-1013-4

Newman, E. (2014). Statebuilding, Upheaval and Civil War. In E. Newman & K. DeRouen (Eds.), *Routledge handbook of civil wars* (p. 173-185). New York: Routledge.

Newman, E., & DeRouen, K. (2014). Introduction. In E. Newman & K. DeRouen (Eds.), *Routledge handbook of civil wars* (p. 1–10). New York: Routledge.

North, D., Wallis, J., & Weingast, B. (2009). *Violence and Social Orders: A conceptual framework for interpreting recorded human history.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511575839

Olson, M. (1993). Dictatorship, Democracy, and Development. *The American Political Science Review*, *87*(3), 567–576. doi: 10.2307/2938736

Organski, A., & Kugler, J. (1980). *The War Ledger.* Chicago, IL: University of Chicago Press.

Osborne, J. (2014). *Best Practices in Exploratory Factor Analysis.* Scotts Valley, CA.

Østby, G. (2008). Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research*, *45*(2), 143-162. doi: 10.1177/0022343307087169

Pettersson, T., Högbladh, S., & Öberg, M. (2019). Organized Violence, 1989-2018 and Peace Agreements. *Journal of Peace Research*, *56*(4), 589-603. doi: 10.1177/0022343319856046

Pettersson, T., & Öberg, M. (2020). Organized violence, 1989-2019. *Journal of Peace Research*, *57*(4). doi: 10.1177/0022343320934986

Pinker, S. (2011). *The Better Angels of Our Nature: Why Violence Has Declined.* New York: Viking Books.

Poteete, A., Janssen, M., & Ostrom, E. (2010). *Working Together: Collective Action, the Commons, and Multiple Methods in Practice.* Princeton University Press.

Prichard, W., Cobham, A., & Goodall, A. (2014). *The ICTD Government Revenue Dataset* (Working Paper No. 19). Institute of Development Studies.

Reno, W. (2000). Clandestine Economies, Violence and States in Africa. *Journal of International Affairs*, *53*(2), 433-459. doi: 10.2307/24357760

Ricciuti, R., & Rossignoli, D. (2017). Bridging Economics and International Relations to Understand State Capacity and War in Sub-Saharan Africa. *Peace Economics, Peace Science and Public Policy*, *23*(4), 1-8. doi: 10.1515/peps-2017-0026

Ross, M. (2015). What Have We Learned About the Resource Curse? *Annual Review of Political Science*, *18*, 239-259. doi: 10.1146/annurev-polisci-052213-040359

Rossignoli, D. (2016). Democracy, State Capacity and Civil Wars: A New Perspective. *Peace Economics, Peace Science and Public Policy*, *22*(4), 427-437. doi: 10.1515/peps-2016-0029

Rotberg, R. (2003). Failed states, collapsed states, weak states: Causes and indicators. In R. Rotberg (Ed.), *When states fail: Causes and consequences* (p. 1–25).

Rubin, D. (1976). Inference and Missing Data. *Biometrika*, *63*(3), 581-592. doi: 10.2307/2335739

Salehyan, I., & Gleditsch, K. (2006). Refugees and the Spread of Civil War. *International Organization*, *60*(2), 335– 366. doi: 10.1017/S0020818306060103

Sanín, F. (2011). Evaluating State Performance: A Critical View of State Failure and Fragility Indexes. *The European Journal of Development Research*, *23*(1), 20-42. doi: 10.1057/ejdr.2010.53

Savoia, A., & Sen, K. (2015). Measurement, evolution, determinants, and consequences of state capacity: a review of recent research. *Journal of Economic Surveys*, *29*(3), 441-458. doi: 10.1111/joes.12065

Schubiger, L., & Zelina, M. (2017). Ideology in Armed Groups. *PS: Political Science & Politics*, *50*(4), 948–952. doi: 10.1017/S1049096517001056

Seawright, J., & Collier, D. (2014). Rival Strategies of Validation: Tools for Evaluating Measures of Democracy. *Comparative Political Studies*, *47*(1), 111-138. doi: 10.1177/0010414013489098

Singer, D., Bremer, S., & Stuckey, J. (1972). Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In B. Russett (Ed.), *Peace, war, and numbers* (p. 19-48). Beverly Hills: Sage.

Skocpol, T. (1979). *States and Social Revolutions: A comparative analysis of France, Russia and China*. Cambridge: Cambridge University Press.

Skocpol, T. (1985). Bringing the State Back In: Strategies of Analysis in Current Research. In P. Evans, D. Rueschemeyer, & T. Skocpol (Eds.), *Bringing the state back in* (p. 3–38). New York and Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511628283.002

Smith, A. (2010). *Nationalism: Theory, ideology, history*. Cambridge: Polity Press.

Smith, R. (2017). Military Expenditure Data: Theoretical and Empirical Considerations. *Defence and Peace Economics*, *28*(4), 422-428. doi: 10.1080/10242694.2016.1245823

Snyder, G. (1971). "Prisoner's Dilemma" and "Chicken" Models in International Politics. *International Studies Quarterly*, *15*(1), 66-103. doi: 10.2307/3013593

Snyder, R., & Bhavnani, R. (2005). Diamonds, Blood, and Taxes: A Revenue-centered Framework for Explaining Political Order. *Journal of Conflict Resolution*, *49*(4), 563-597. doi: 10.1177/0022002705277796

Sobek, D. (2010). Masters of their domains: The role of state capacity in civil wars. *Journal of Peace Research*, *47*(3), 267–271. doi: 10.1177/0022343310362295

Soifer, H. (2008). State Infrastructural Power: Approaches to Conceptualization and Measurement. *Studies in Comparative International Development*, *43*, 231-251. doi: 0.1007/s12116-008-9028-6

Soifer, H. (2009, May). *The Redistributive Threat: State Power and the Effect of Inequality on Democracy* (Working Paper No. 93). University of Manchester: Brooks World Poverty Institute.

Soifer, H. (2012). Measuring State Capacity in Contemporary Latin America. *Revista de Ciencia Política*, *32*(3), 585-598. Retrieved from `https://www.redalyc.org/articulo.oa?id=32425402004`

Soifer, H. (2013). State Power and the Economic Origins of Democracy. *Studies in Comparative International Development*, *48*, 1–22. doi: 10.1007/s12116-012-9122-7

Soifer, H., & Vom Hau, M. (2008). Unpacking the Strength of the State: The

Utility of State Infrastructural Power. *Studies in Comparative International Development*, *43*(3), 219-230. doi: 10.1007/s12116-008-9030-z

Starr, H. (1978). "Opportunity" and "Willingness" as Ordering Concepts in the Study of War. *International Interactions*, *4*, 363-378. doi: 10.1080/03050627808434499

Stewart, F. (2008). *Horizontal inequalities and conflict: Understanding group violence in multiethnic societies*. Basingstoke: Palgrave Macmillan.

Stiftung, B. (2020). Transformation Index of the Bertelsmann Stiftung 2020: Codebook for Country Assessments. Retrieved from `https://www.bti-project.org/content/en/downloads/codebooks/BTI_2020_Codebook.pdf`

Strand, H., Rustad, S., Nygård, H., & Hegre, H. (2020). Trends in armed conflict, 1946-2019. *PRIO: Conflict Trends*, *8*. Retrieved from `https://www.prio.org/Publications/Publication/?x=12532`

Sundberg, R., & Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, *50*(4), 523-532. doi: 10.1177/0022343313484347

Taydas, Z., Peksen, D., & James, P. (2010). Why Do Civil Wars Occur? Understanding the Importance of Institutional Quality. *Civil Wars*, *12*(3), 195-217. doi: 10.1080/13698249.2010.509544

Teorell, J., Sundström, A., Holmberg, S., Rothstein, B., Alvarado, N., & Dalli, C. (2021). The Quality of Government Standard Dataset, version Jan21. *University of Gothenburg: The Quality of Government Institute*. Retrieved from `http://www.qog.pol.gu.sedoi`

Thies, C. (2010). Of Rulers, Rebels, and Revenue: State capacity, Civil War Onset, and Primary Commodities. *Journal of peace research*, *47*(3), 321-332. doi: 10.1177/0022343310361841

Thies, C. (2015). State fiscal capacity and state failure. *European Political Science Review: EPSR*, *7*(2), 167-187. doi: 10.1017/S175577391400037X

Thyne, C. (2006). ABC's, 123's, and the golden rule: The pacifying effect of education on civil war, 1980-1999. *International Studies Quarterly*, *50*(4), 733-754.

Tilly, C. (1978). *From mobilization to revolution*. New York: McGraw-Hill.

Tilly, C. (1985). War Making and State Making as Organized Crime. In P. Evans, D. Rueschemeyer, & T. Skocpol (Eds.), *Bringing the state back in* (p. 169–190). doi: 10.1017/CBO9780511628283.008

Tilly, C. (1993). *Coercion, Capital and European States, A.D. 990-1992*. Cambridge: Wiley-Blackwell.

Tilly, C. (2003). Inequality, Democratization, and De-Democratizationn. *Sociological theory*, *21*(1), 37-43. doi: 10.1111/1467-9558.00174

Treier, S., & Jackman, S. (2008). Democracy as a Latent Variable. *American Journal of Political Science*, *52*(1), 201-217. doi: 10.1111/j.1540-5907.2007.00308.x

Vaccaro, A. (2020). *Measures of State Capacity: Same Same, but Different?* (Working Paper No. 9). Department of Political Science, University of Gothenburg: The Quality of Government Institute. Retrieved from `https://www.gu.se/nyheter/nytt-qog-working-paper-measures-of-state-capacity-same-same-but-different`

Van Buuren, S. (2018). *Flexible Imputation of Missing Data*. Boca Raton: Chapman and Hall/CRC Press.

Verweijen, J. (2018). Soldiers Without an Army? Patronage Networks and Cohesion in the Armed Forces of the DR Congo. *Armed Forces & Society*, *44*(4), 626-646. doi: 10.1177/0095327X17740096

Vink, G., Frank, L., Pannekoek, J., & Van Buuren, S. (2014). Predictive Mean Matching Imputation of Semicontinuous Variables. *Statistica Neerlandica*, *68*(1), 61-90. doi: 10.1111/stan.12023

Von Einsiedel, S., Bosetti, L., Cockayne, J., Salih, C., & Wan, W. (2017). Civil War Trends and the Changing Nature of Armed Conflict. *United Nations University Centre for Policy Research, Occasional Paper 10*. Retrieved from `https://collections.unu.edu/eserv/UNU:6156/Civil_war_trends_UPDATED.pdf`

Wai, Z. (2012). Neo-patrimonialism and the Discourse of State Failure in Africa. *Review of African Political Economy*, *39*(131), 27-43. doi: 10.1080/03056244.2012.658719

Walter, B. (2006). Building Reputation: Why Governments Fight Some Separatists

but Not Others. *American Journal of Political Science*, *50*(2), 313–330. doi: 10.1111/j.1540-5907.2006.00186.x

Walter, B. (2017). The Extremist's Advantage in Civil Wars. *International Security*, *42*(2), 7–39. doi: 10.1162/ISEC_a_00292

Ward, M., Greenhill, B., & Bakke, K. (2010). The Perils of Policy by p-value: Predicting Civil Conflicts. *Journal of Peace Research*, *47*(4), 363–375. doi: 10.1177/0022343309356491

Weber, M. (1918). *From Max Weber: Essays in Sociology*. Oxford: Oxford University Press.

Wegenast, T., & Basedau, M. (2014). Ethnic Fractionalization, Natural Resources and Armed Conflict. *Conflict Management and Peace Science*, *31*(4), 432–457. doi: 10.1177/0738894213508692

Wilkinson, C. (2014). Development in Kyrgyzstan: Failed State or Failed State-building? In A. Ware (Ed.), *Development in difficult sociopolitical contexts* (p. 137-162).

Williams, M. (2020). Beyond State Capacity: Bureaucratic Performance, Policy Implementation and Reform. *Journal of Institutional Economics*, *17*(2), 1–19. doi: 10.1017/S1744137420000478

Wucherpfennig, J., Metternich, N., Cederman, L.-E., & Gleditsch, K. (2012). Ethnicity, the state, and the duration of civil war. *World Politics*, *64*(1), 79-115. doi: 10.1017/S004388711100030X

Yong, A., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysiss. *Tutorials in quantitative methods for psychology*, *9*(2), 79-94. doi: 10.20982/tqmp.09.2.p079

Zakaria, F. (1997). The Rise of Illiberal Democracy. *Foreign Affairs*, *76*(6), 22-43. doi: 10.2307/20048274

Zakaria, F. (2007). *The Future of Freedom: Illiberal Democracy at Home and Abroad*. New York: W.W. Norton & Company.

Zhang, Z. (2016). Missing Data Imputation: Focusing on Single Imputation. *Ann Transl Med*, *4*(1). doi: 10.3978/j.issn.2305-5839.2015.12.38

Appendix A

Descriptives

TABLE A.1: Data coverage

Countries[a]	*Region*[b]	*Years*[c]		*Indicators*[d]		
		min.	*max.*	*min.*	*max.*	*median*
Afghanistan	South Asia	1991	2019	7	15	10
Albania	E. Europe & post-USSR	1991	2019	13	19	17
Algeria	N. Africa & the Middle East	1991	2019	10	16	14
Angola	Sub-Saharan Africa	1991	2019	10	18	17
Argentina	Latin America	1991	2019	13	19	17
Armenia	E. Europe & post-USSR	1991	2019	9	18	15
Australia	W. Europe & N. America	1991	2019	12	19	18
Austria	W. Europe & N. America	1991	2019	12	19	18
Azerbaijan	E. Europe & post-USSR	1991	2019	7	18	16
Bahrain	N. Africa & the Middle East	1991	2019	9	17	15
Bangladesh	South Asia	1991	2019	10	18	17
Belarus	E. Europe & post-USSR	1991	2019	7	19	16
Belgium	W. Europe & N. America	1991	2019	12	19	18
Benin	Sub-Saharan Africa	1991	2019	7	13	11
Bhutan	South Asia	1991	2019	5	11	11
Bolivia	Latin America	1991	2019	10	19	17
Bosnia-Herzegovina	E. Europe & post-USSR	1992	2019	7	16	13
Botswana	Sub-Saharan Africa	1991	2019	13	18	17
Brazil	Latin America	1991	2019	13	20	16
Bulgaria	E. Europe & post-USSR	1991	2019	13	20	18
Burkina Faso	Sub-Saharan Africa	1991	2019	13	18	16
Burundi	Sub-Saharan Africa	1991	2019	7	14	13
Cambodia	South-East Asia	1991	2019	9	15	13
Cameroon	Sub-Saharan Africa	1991	2019	10	18	16
Canada	W. Europe & N. America	1991	2019	12	19	18
Cape Verde	Sub-Saharan Africa	1991	2019	7	16	11

Continuation of Table A.1

Central African Republic	Sub-Saharan Africa	1991	2019	7	15	12
Chad	Sub-Saharan Africa	1991	2019	7	13	12
Chile	Latin America	1991	2019	13	20	18
China	East Asia	1991	2019	10	19	15
Colombia	Latin America	1991	2019	13	19	16
Comoros	Sub-Saharan Africa	1991	2019	5	9	8
Congo, DR	Sub-Saharan Africa	1991	2019	10	16	15
Congo, Rep	Sub-Saharan Africa	1991	2019	10	18	16
Costa Rica	Latin America	1991	2019	13	20	18
Côte d'Ivoire	Sub-Saharan Africa	1991	2019	13	19	15
Croatia	E. Europe & post-USSR	1992	2019	12	18	16
Cuba	Latin America	1991	2019	7	15	13
Cyprus	N. Africa & the Middle East	1991	2019	12	18	17
Czech Republic	E. Europe & post-USSR	1993	2019	12	18	17
Denmark	W. Europe & N. America	1991	2019	12	18	17
Djibouti	Sub-Saharan Africa	1991	2019	5	13	12
Dominican Republic	Latin America	1991	2019	13	20	18
Ecuador	Latin America	1991	2019	10	17	15
Egypt	N. Africa & the Middle East	1991	2019	10	19	18
El Salvador	Latin America	1991	2019	13	20	18
Equatorial Guinea	Sub-Saharan Africa	1991	2019	8	15	12
Eritrea	Sub-Saharan Africa	1993	2019	5	12	9
Estonia	E. Europe & post-USSR	1992	2019	11	18	17
Eswatini	Sub-Saharan Africa	1991	2019	7	12	12
Ethiopia	Sub-Saharan Africa	1991	2019	13	19	18
Fiji	The Pacific	1991	2019	10	15	14
Finland	W. Europe & N. America	1991	2019	12	19	18
France	W. Europe & N. America	1991	2019	12	19	18
Gabon	Sub-Saharan Africa	1991	2019	12	18	14
Georgia	E. Europe & post-USSR	1991	2019	6	16	14
Germany	W. Europe & N. America	1991	2019	12	19	18
Ghana	Sub-Saharan Africa	1991	2019	13	17	16
Greece	W. Europe & N. America	1991	2019	12	19	18
Guatemala	Latin America	1991	2019	13	20	18
Guinea	Sub-Saharan Africa	1991	2019	10	17	15
Guinea-Bissau	Sub-Saharan Africa	1991	2019	10	16	15
Guyana	The Caribbean	1991	2019	10	17	15
Haiti	Latin America	1991	2019	10	16	15
Honduras	Latin America	1991	2019	10	20	16
Hungary	E. Europe & post-USSR	1991	2019	13	20	18
India	South Asia	1991	2019	10	19	18
Indonesia	South-East Asia	1991	2019	13	20	18
Iran	N. Africa & the Middle East	1991	2019	10	19	18
Iraq	N. Africa & the Middle East	1991	2019	11	17	14
Ireland	W. Europe & N. America	1991	2019	12	18	17

Continuation of Table A.1

Israel	N. Africa & the Middle East	1991	2019	12	17	16
Italy	W. Europe & N. America	1991	2019	12	19	18
Jamaica	The Caribbean	1991	2019	13	18	17
Japan	East Asia	1991	2019	9	17	14
Jordan	N. Africa & the Middle East	1991	2019	13	19	17
Kazakhstan	E. Europe & post-USSR	1991	2019	7	19	16
Kenya	Sub-Saharan Africa	1991	2019	13	19	15
Kuwait	N. Africa & the Middle East	1991	2019	9	17	15
Kyrgyzstan	E. Europe & post-USSR	1991	2019	7	15	13
Laos	South-East Asia	1991	2019	5	13	12
Latvia	E. Europe & post-USSR	1992	2019	8	18	17
Lebanon	N. Africa & the Middle East	1991	2019	13	19	17
Lesotho	Sub-Saharan Africa	1991	2019	10	14	12
Liberia	Sub-Saharan Africa	1991	2019	10	17	14
Libya	N. Africa & the Middle East	1991	2019	8	17	15
Lithuania	E. Europe & post-USSR	1992	2019	10	18	17
Luxembourg	W. Europe & N. America	1991	2019	12	17	16
Madagascar	Sub-Saharan Africa	1991	2019	13	19	18
Malawi	Sub-Saharan Africa	1991	2019	13	18	15
Malaysia	South-East Asia	1991	2019	13	18	17
Maldives	South Asia	1991	2019	5	11	10
Mali	Sub-Saharan Africa	1991	2019	12	18	16
Malta	W. Europe & N. America	1991	2019	12	17	16
Mauritania	Sub-Saharan Africa	1991	2019	7	13	12
Mauritius	Sub-Saharan Africa	1991	2019	10	16	14
Mexico	Latin America	1991	2019	13	20	18
Moldova	E. Europe & post-USSR	1992	2019	6	19	16
Mongolia	East Asia	1991	2019	10	18	17
Montenegro	E. Europe & post-USSR	2006	2019	7	15	14
Morocco	N. Africa & the Middle East	1991	2019	13	20	18
Mozambique	Sub-Saharan Africa	1991	2019	13	18	16
Myanmar	South-East Asia	1991	2019	11	19	18
Namibia	Sub-Saharan Africa	1991	2019	10	18	17
Nepal	South Asia	1991	2019	10	16	15
Netherlands	W. Europe & N. America	1991	2019	12	19	18
New Zealand	W. Europe & N. America	1991	2019	12	19	17
Nicaragua	Latin America	1991	2019	13	19	17
Niger	Sub-Saharan Africa	1991	2019	10	16	15
Nigeria	Sub-Saharan Africa	1991	2019	10	16	15
North Macedonia	E. Europe & post-USSR	1991	2019	6	16	13
Norway	W. Europe & N. America	1991	2019	12	19	18
Oman	N. Africa & the Middle East	1991	2019	9	14	14
Pakistan	South Asia	1991	2019	10	16	15
Panama	Latin America	1991	2019	10	18	15
Papua New Guinea	The Pacific	1991	2019	13	18	15

Continuation of Table A.1

Paraguay	Latin America	1991	2019	13	20	16
Peru	Latin America	1991	2019	13	20	16
Philippines	South-East Asia	1991	2019	13	19	17
Poland	E. Europe & post-USSR	1991	2019	13	20	18
Portugal	W. Europe & N. America	1991	2019	12	19	18
Qatar	N. Africa & the Middle East	1991	2019	7	15	11
Romania	E. Europe & post-USSR	1991	2019	13	19	18
Russia	E. Europe & post-USSR	1991	2019	7	20	18
Rwanda	Sub-Saharan Africa	1991	2019	10	15	13
Saudi Arabia	N. Africa & the Middle East	1991	2019	12	18	16
Senegal	Sub-Saharan Africa	1991	2019	10	18	14
Serbia	E. Europe & post-USSR	1995	2019	9	20	15
Sierra Leone	Sub-Saharan Africa	1991	2019	10	17	15
Singapore	South-East Asia	1991	2019	12	18	17
Slovakia	E. Europe & post-USSR	1993	2019	13	19	17
Slovenia	E. Europe & post-USSR	1992	2019	12	18	17
Somalia	Sub-Saharan Africa	1991	2019	10	13	11
South Africa	Sub-Saharan Africa	1991	2019	13	18	17
South Korea	East Asia	1991	2019	9	19	18
South Sudan	Sub-Saharan Africa	2011	2019	7	11	10
Spain	W. Europe & N. America	1991	2019	12	17	16
Sri Lanka	South Asia	1991	2019	13	18	17
Sudan	Sub-Saharan Africa	1991	2019	10	18	14
Suriname	The Caribbean	1991	2019	8	13	12
Sweden	W. Europe & N. America	1991	2019	12	19	18
Switzerland	W. Europe & N. America	1991	2019	12	19	18
Syria	N. Africa & the Middle East	1991	2019	8	16	13
Tajikistan	E. Europe & post-USSR	1991	2019	5	15	12
Tanzania	Sub-Saharan Africa	1991	2019	10	19	14
Thailand	South-East Asia	1991	2019	13	20	18
The Gambia	Sub-Saharan Africa	1991	2019	10	16	15
Timor-Leste	South-East Asia	2002	2019	9	16	13
Togo	Sub-Saharan Africa	1991	2019	13	18	15
Trinidad & Tobago	The Caribbean	1991	2019	10	19	17
Tunisia	N. Africa & the Middle East	1991	2019	10	19	18
Turkey	N. Africa & the Middle East	1991	2019	13	20	17
Turkmenistan	E. Europe & post-USSR	1991	2019	5	12	10
Uganda	Sub-Saharan Africa	1991	2019	13	18	14
Ukraine	E. Europe & post-USSR	1991	2019	7	19	17
United Arab Emirates	N. Africa & the Middle East	1991	2019	10	15	13
United Kingdom	W. Europe & N. America	1991	2019	12	18	18
United States of America	W. Europe & N. America	1991	2019	12	19	18
Uruguay	Latin America	1991	2019	13	20	18
Uzbekistan	E. Europe & post-USSR	1991	2019	7	14	12
Venezuela	Latin America	1991	2019	8	17	16

End of Table A.1

Vietnam	South-East Asia	1991	2019	8	17	14
Yemen	N. Africa & the Middle East	1991	2019	8	17	16
Zambia	Sub-Saharan Africa	1991	2019	13	18	16
Zimbabwe	Sub-Saharan Africa	1991	2019	10	18	15

[a] N=165

[b] Politico-geographic regions, taken from Teorell et al. (2021)

[c] 29 possible years (1991-2019)

[d] The minimum number of available indicators per country-year is 5, and the maximum is 20. The maximum number of country-years is 4722. In 96% of country-years, at least 10 indicators are available, and the median number of indicators per country-year is 16.

Figure A.1: Qualitative description of indicators

`milper_1000capita (IISS & CoW)` – Armed forces personnel per 1000 capita. Active duty military personnel, including paramilitary forces if the training, organization, equipment, and control suggest they may be used to support or replace regular military forces. Data excludes personnel not on active duty. Where IISS reports missing information, I have added data from CoW (a total of 73 country-years).

`milex_prcgdp (SIPRI)` – Military expenditures as % of GDP. An input measure of the share of military expenditure in GDP, indicating the priority attached to the military and how much of available output is devoted to defence.

`milex_capita (SIPRI)` – Military expenditures per capita. Includes all current and capital expenditures on the armed forces, including peacekeeping forces; defense ministries and other government agencies engaged in defense projects; paramilitary forces, if these are judged to be trained and equipped for military operations; and military space activ- ities. Figures are in US$m at constant 2018 prices and exchange rates.

`statautterr (V-Dem v.11.1)` – State authority over territory. Over what percentage (%) of the territory does the state have effective control? Estimates the size of the territory that the state has effective control over, as a percentage of the total territory that is officially part of the country. By "officially part of the country", V-Dem refers to international law. In the V-Dem dataset, the variable is called "v2svstterr".

`poloff (UNODC)` – Police personnel per 100,000 capita. Police personnel in public agencies as at 31 December whose principal functions are the prevention, detection and investigation of crime and the apprehension of alleged offenders.

`icrg_laworder (ICRG)` – Law and order. The measure form a single component, but its two elements are assessed separately, with each element being scored from zero to three points. To assess the "Law" element, the strength and impartiality of the legal system are considered, while the "Order" element is an assessment of popular observance of the law. The index is measured on 0-6 scale, ranging from low (0) to high (6) levels of law and order.

`statehist (Bockstette et al., 2002 & Hanson & Sigman, 2020)` – State Antiquity Index. Measures the degree to which each of the present-day countries was the site of nation-states, kingdoms, or empires every half-a-century since the first century. Attempting to capture the state's level of institutionalization or presence in the territory. The measure is extended by Hanson and Sigman (2020).

`icrg_bq (ICRG)` – Bureaucratic quality. High points are given to countries where the bureaucracy has the strength and expertise to govern without drastic changes in policy or interruptions in government services. In these countries, the bureaucracy tends to be somewhat autonomous from political pressure and to have an established mechanism for recruitment and training. Measured on a 0-4 scale, ranging from low (0) to high (4) bureaucratic quality.

`icrg_corruption (ICRG)` – Corruption within the political system. The corruption index ranges from 0 to 6, where 0 corresponds to the highest possible level of corruption and 6 to the lowest possible level of corruption. It captures the following spheres of illegal activity: actual or potential corruption in the form of excessive patronage, nepotism, job reservations, 'favor-for-favors', secret party funding, and suspiciously close ties between politics and business.

`pubadmimpartial (V-Dem v.11.1)` – Rigorous and impartial public administration. Are public officials rigorous and impartial in the performance of their duties? Focuses on the extent to which public officials generally abide by the law and treat like cases alike, or conversely, the extent to which public administration is characterized by arbitrariness and biases (i.e., nepotism, cronyism, or discrimination). The question covers the public officials that handle the cases of ordinary people. In the V-Dem dataset, the variable is called "v2clrspct".

`critrecadm (V-Dem v.11.1)` – Criteria for appointment decisions in the state administration. To what extent are appointment decisions in the state administration based on personal and political connections, as opposed to skills and merit? Appointment decisions include hiring, firing and promotion in the state administration. The question refers to the typical de facto (rather than de jure) situation obtaining in the state administration, excluding the armed forces. In the V-Dem dataset, the variable is called "v2stcritrecadm".

`statcap (World Bank)` – Statistical capacity. Captures a nation's ability to collect, analyze, and disseminate high-quality data about its population and economy. It is based on a diagnostic framework assessing the following areas: methodology; data sources; and periodicity and timeliness. Countries are scored against 25 criteria in these areas, using publicly available information and/or country input. The overall Statistical Capacity score is then calculated as a simple average of all three area scores on a scale of 0-100.

`infcap (Brambor et al., 2020)` – Information capacity. An aggregate index (0-1) based on five indicators of information-gathering and information- organizing activities: the introduction and regular implementation of a national census, the introduction of civil registers the introduction of population registers, the establishment of a permanent state agency tasked with processing statistical in- formation about a country's territory and population, and the regular publication of statistical yearbooks.

`censusfreq (Hanson & Sigman, 2020 & United States Census Bureau)` – Census frequency. Hanson and Sigman (2020) derive a measure of census frequency calculated with data on country censuses provided by the U.S. Census Bureau. They have further annualized this measure by looking forward and backward in time from a given year to find the nearest censuses. The longer the gaps between censuses, the lower the Census Frequency measure.

`taxrev_prcgdp (ICTD & World Bank)` – Tax revenue as % of GDP. Tax revenue refers to compulsory transfers to the central government for public purposes. Certain compulsory transfers such as fines, penalties, and most social security contributions are excluded.

`incometax_prctax (ICTD & World Bank)` – Taxes on income as % of total tax revenue. Taxes on income, profits, and capital gains are levied on the actual or presumptive net income of individuals, on the profits of corporations and enterprises, and on capital gains, whether realized or not, on land, securities, and other assets. Intragovernmental payments are eliminated in consolidation.

`tradetax_prctax (ICTD & World Bank)` – Taxes on international trade as % of total tax revenue. Taxes on international trade include import duties, export duties, profits of export or import monopolies, exchange profits, and exchange taxes.

`relpolcap (Fisunoglu et al., 2011)` – Relative political capacity. Assesses the amount of tax collected relative to an estimated expected amount of revenue. Also, it compares a state's efficiency at extracting resources compared with other states with similar endowments.

`resourcerev_prcgdp (ICTD)` – Natural resource revenue as % of GDP. The indicator is calculated using ICTD's "total non tax revenue" as the difference between "total revenue" and "total non-resource tax revenue". This variable will thus include all types of non-tax revenue, including resource revenue, but acts as a useful proxy for resource revenue because resource revenue explains the vast majority of the variation in "total non tax revenue" across countries.

`fisccap (V-Dem v.11.1))` – Fiscal capacity. Captures the extent to which the state is able to fund itself through taxes that are of greater administrative complexity. In the V-Dem dataset, the variable is called "v2stfisccap".

Appendix B

Missingness

FIGURE B.1: Missingness map

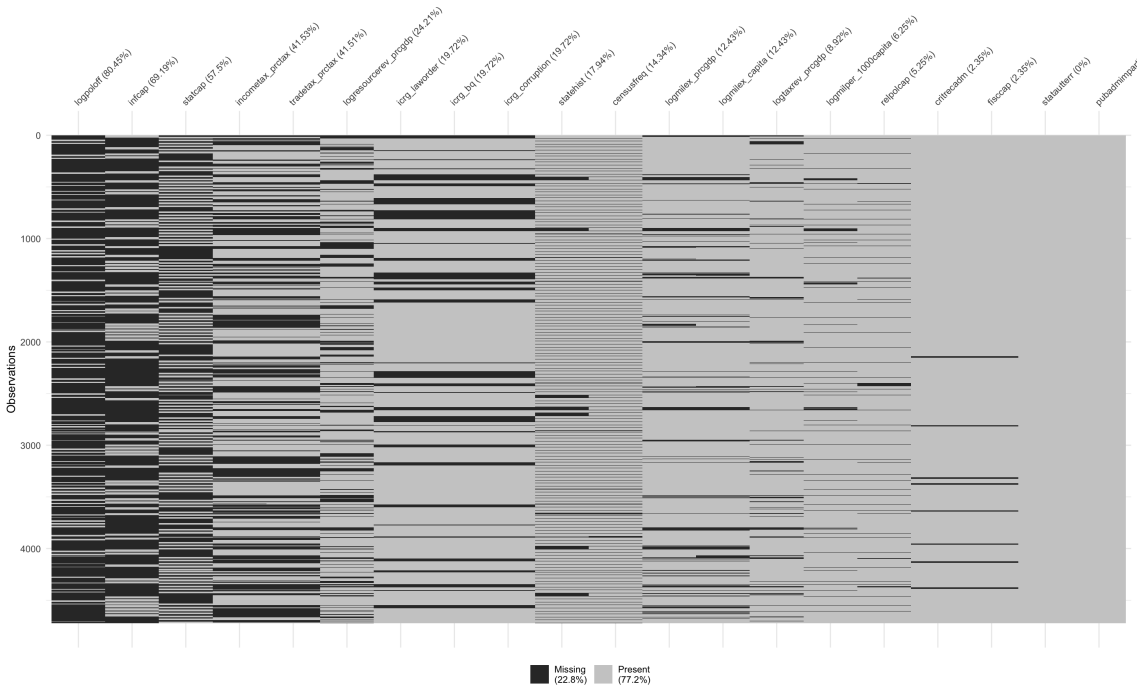


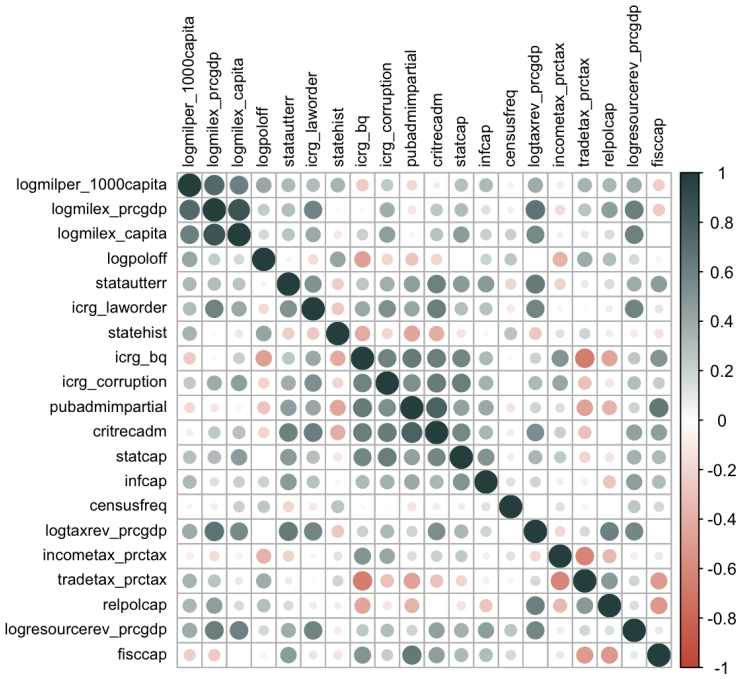
Table B.1: Comparison of imputation models

	Summary statistics by imputation model: mean (std.error)										
	0[a]	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
milper_1000capita	6.511878	6.514327	6.48265	6.458957	6.479429	6.509283	6.518891	6.472948	6.465679	6.482855	6.481458
	(.1112)	(.1069)	(.1075)	(.1070)	(.1069)	(.1073)	(.1093)	(.1072)	(.1068)	(.1071)	(.1074)
milex_prcgdp	2.247591	2.322671	2.322131	2.354702	2.311804	2.330262	2.36528	2.3171	2.33859	2.339498	2.343173
	(.0447)	(.0419)	(.0411)	(.0429)	(.0416)	(.0419)	(.0429)	(.0422)	(.0430)	(.0420)	(.0418)
milex_capita	263.4787	266.6731	266.053	269.6642	263.5004	263.9378	266.5922	262.9831	269.0244	261.9501	267.3305
	(8.127)	(7.7142)	(7.7313)	(7.8534)	(7.6810)	(7.6040)	(7.7168)	(7.5833)	(7.8668)	(7.6124)	(7.8647)
poloff	318.6479	287.5235	266.8621	278.1214	267.6866	282.9087	268.0844	284.6629	271.0339	287.9963	288.1134
	(5.1610)	(3.0943)	(2.8586)	(2.9956)	(2.8282)	(2.9210)	(2.7961)	(2.8327)	(2.7640)	(3.0318)	(2.9433)
statautterr	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599	91.23599
	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)	(.1593)
icrg_laworder	3.720549	3.599513	3.6036	3.600932	3.609466	3.612579	3.609932	3.60809	3.611563	3.605845	3.605421
	(.0223)	(.0196)	(.0196)	(.0195)	(.0195)	(.0195)	(.0195)	(.0196)	(.0196)	(.0195)	(.0196)
statehist	.6994839	.7019208	.7028484	.7041148	.7030093	.7048687	.7030072	.7033122	.703274	.7030771	.7050678
	(.0026)	(.0024)	(.0024)	(.0023)	(.0024)	(.0024)	(.0024)	(.0024)	(.0024)	(.0024)	(.0024)
icrg_bq	2.162833	2.023973	2.024312	2.02698	2.03363	2.017323	2.025074	2.019123	2.021432	2.007391	2.020267
	(.0182)	(.0162)	(.0161)	(.0161)	(.0160)	(.0162)	(.0161)	(.0161)	(.0161)	(.0163)	(.0162)
icrg_corruption	3.183566	3.306078	3.301906	3.304828	3.3	3.311542	3.304807	3.297861	3.293986	3.313236	3.299767
	(.0204)	(.0177)	(.0177)	(.0177)	(.0176)	(.0176)	(.0176)	(.0176)	(.0177)	(.0177)	(.0177)
pubadmimpartial	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235	.3745235
	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)	(.0217)
critrecarm	.4284906	.4279903	.429928	.4296442	.4292397	.4294028	.4268996	.4262685	.4284689	.4271135	.4264019
	(.0159)	(.0156)	(.0156)	(.0156)	(.0156)	(.0156)	(.0156)	(.0156)	(.0156)	(.0156)	(.0157)
statcap	67.40386	67.02242	66.12859	66.88578	67.32262	67.54308	67.08984	67.03708	66.94084	66.32354	66.75947
	(.3648)	(.2575)	(.2596)	(.2558)	(.2619)	(.2575)	(.2562)	(.2595)	(.2512)	(.2626)	(.2554)
infcap	.6487285	.593291	.5907836	.5841847	.5844473	.5883079	.5889496	.5866201	.587842	.5893139	.5837357
	(.0039)	(.0021)	(.0022)	(.0022)	(.0022)	(.0022)	(.0022)	(.0022)	(.0022)	(.0022)	(.00222)
censusfreq	.9707686	.9702531	.9693585	.9693998	.9729164	.9715994	.9727539	.9714846	.972068	.9722807	.971994
	(.0059)	(.0054)	(.0054)	(.0055)	(.0055)	(.0055)	(.0055)	(.0054)	(.0055)	(.0055)	(.0054)
taxrev_prcgdp	16.67933	16.35418	16.2852	16.28511	16.32833	16.33999	16.31281	16.32483	16.26337	16.33557	16.31932
	(.1345)	(.1310)	(.1296)	(.1284)	(.1311)	(.1285)	(.1312)	(.1307)	(.1304)	(.1292)	(.1308)
incometax_prctax	36.0108	33.6093	33.4693	33.47759	33.83159	33.81976	33.76493	34.2211	33.80822	33.2887	33.95812
	(.3258)	(.2418)	(.2436)	(.2423)	(.2416)	(.2392)	(.2454)	(.2417)	(.2430)	(.2460)	(.2442)
tradetax_prctax	11.58185	16.59933	16.75027	16.85522	16.89875	16.71088	17.01487	16.6671	16.62897	17.26445	17.0694
	(.2671)	(.2439)	(.2512)	(.2500)	(.2509)	(.2478)	(.2497)	(.2459)	(.2449)	(.2553)	(.2549)
relpolcap	.9946371	1.001652	.9989436	.9980631	.9968173	.9973253	.9980028	.9976781	.9950768	.9989294	.9981551
	(.0058)	(.0057)	(.0057)	(.0057)	(.0057)	(.0057)	(.0058)	(.0058)	(.0057)	(.0057)	(.0058)
resourcerev_prcgdp	6.840471	7.66139	7.777811	8.018951	7.807544	7.798883	7.59956	7.70108	7.845883	7.730958	7.868802
	(.1487)	(.1495)	(.1564)	(.1604)	(.1561)	(.1566)	(.1486)	(.1535)	(.1595)	(.1538)	(.1566)
fisccap	.9136586	.9082317	.9089094	.905972	.9085282	.9092524	.906474	.905432	.9070796	.9066942	.9076662
	(.0180)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)	(.0177)

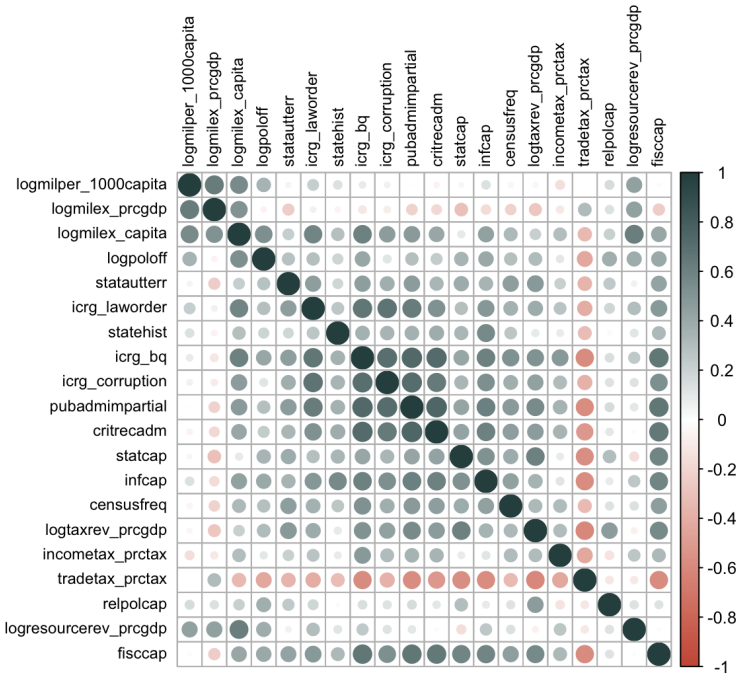
Notes: (a) model 0 equals the observed data

FIGURE B.2: Correlation plots before and after imputation

(A) Before imputation



(B) After imputation



Appendix C

Factor Analysis

FIGURE C.1: Scree plot for determining the number of factors

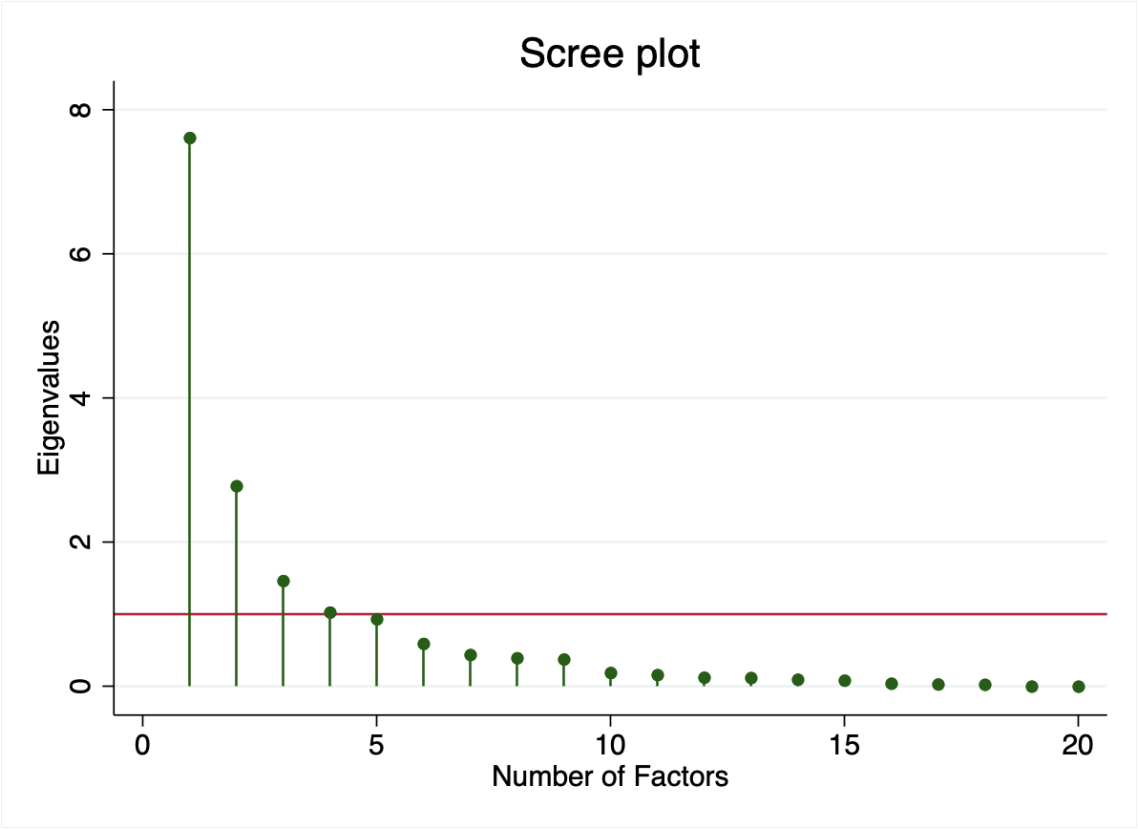


FIGURE C.2: Parallel analysis for determining the number of factors

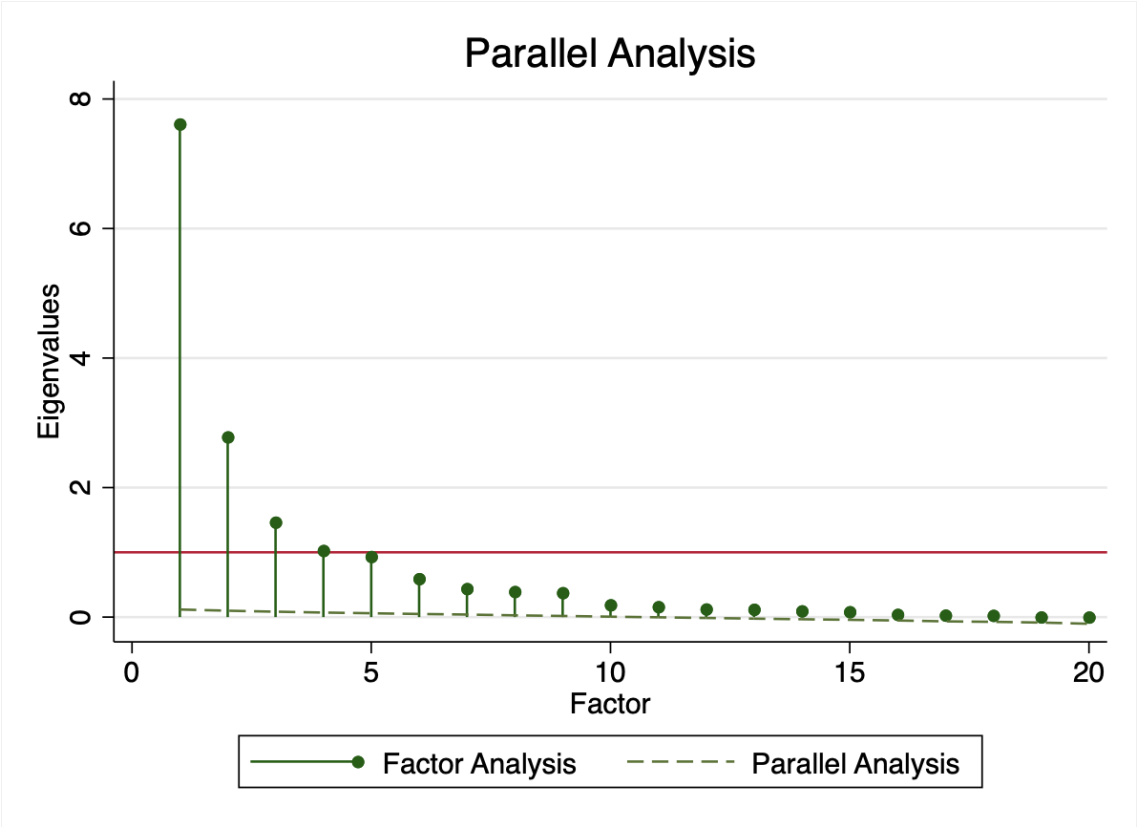


TABLE C.1: Principal factor analysis full rank orderings, 2019

Rank orderings	*Factor 1*	*Factor 2*	*Factor 3*
1	Denmark	Serbia	Libya
2	Germany	Croatia	Saudi Arabia
3	Luxembourg	Bulgaria	Syria
4	Sweden	Kyrgyzstan	Kuwait
5	New Zealand	Belarus	Oman
6	Ireland	Mongolia	Bahrain
7	United Kingdom	North Macedonia	Yemen
8	Norway	Armenia	Russia
9	Switzerland	Moldova	Israel
10	Netherlands	Montenegro	Qatar
11	Australia	Poland	Somalia
12	Singapore	El Salvador	Azerbaijan
13	Finland	Slovenia	Algeria
14	Canada	Georgia	Jordan
15	Belgium	Mauritius	Lebanon
16	Japan	Czech Republic	Singapore
17	United States of America	Albania	Greece
18	Austria	Chile	Iraq
19	Spain	Turkey	United States of America
20	France	Sweden	Turkey
21	Slovenia	Malta	United Arab Emirates
22	Israel	Russia	Armenia
23	South Korea	Cuba	Romania
24	Cyprus	Portugal	Bulgaria
25	Portugal	Mexico	South Korea
26	Czech Republic	Bosnia-Herzegovina	Iran
27	Chile	Romania	Croatia
28	Estonia	Costa Rica	Montenegro
29	Uruguay	Greece	Finland
30	Malta	Ukraine	Turkmenistan
31	India	Thailand	Ukraine
32	Italy	Sri Lanka	Serbia
33	United Arab Emirates	Kazakhstan	France
34	Latvia	Argentina	Poland
35	Malaysia	Vietnam	Norway
36	Indonesia	Brazil	Eritrea
37	Slovakia	Egypt	Ecuador
38	Poland	Latvia	Italy
39	Greece	Uruguay	Australia
40	Argentina	Japan	China
41	Hungary	Jordan	Colombia
42	Lithuania	Colombia	Kazakhstan
43	Croatia	Cyprus	Belarus

Continuation of Table C.1

44	Peru	Netherlands	Cuba
45	Bhutan	Austria	Latvia
46	Kazakhstan	United Kingdom	United Kingdom
47	Botswana	Dominican Republic	Botswana
48	Brazil	South Africa	Portugal
49	Trinidad & Tobago	Finland	Congo, Rep
50	Tunisia	Panama	Netherlands
51	Namibia	Hungary	Uruguay
52	Costa Rica	Azerbaijan	Estonia
53	Mexico	Spain	Cyprus
54	China	Italy	Morocco
55	Mauritius	Belgium	Chile
56	Maldives	New Zealand	New Zealand
57	Bulgaria	South Korea	Sweden
58	Philippines	Denmark	Eswatini
59	Benin	China	Czech Republic
60	Mongolia	Tajikistan	Slovakia
61	Qatar	Togo	Maldives
62	South Africa	Peru	North Macedonia
63	Montenegro	Philippines	Denmark
64	Kenya	Uzbekistan	Lithuania
65	Jamaica	Senegal	Bosnia-Herzegovina
66	Georgia	Israel	Tunisia
67	Colombia	Iran	Luxembourg
68	Albania	Maldives	Spain
69	El Salvador	Ecuador	Gabon
70	Guyana	France	Austria
71	Thailand	Eswatini	Uzbekistan
72	Sri Lanka	Fiji	Cambodia
73	Moldova	Malaysia	Slovenia
74	Vietnam	Laos	Tajikistan
75	The Gambia	Australia	Venezuela
76	Oman	Cambodia	Fiji
77	Serbia	Cape Verde	Sri Lanka
78	Romania	Tunisia	Canada
79	Timor-Leste	Indonesia	Hungary
80	Guatemala	Estonia	Switzerland
81	Ghana	Jamaica	Jamaica
82	Iran	Paraguay	Namibia
83	Bolivia	Morocco	Angola
84	North Macedonia	Bolivia	Belgium
85	Armenia	Slovakia	Germany
86	Ecuador	Honduras	Malta
87	Cuba	Malawi	Mauritania

Continuation of Table C.1

88	Saudi Arabia	Nepal	Brazil
89	Kuwait	Guatemala	Malaysia
90	Morocco	Lesotho	Thailand
91	Belarus	Sudan	Albania
92	Jordan	Germany	Georgia
93	Panama	Lithuania	Dominican Republic
94	Lesotho	The Gambia	Japan
95	Kyrgyzstan	Ghana	Peru
96	Russia	Ireland	Bolivia
97	Fiji	Bangladesh	Pakistan
98	Myanmar	Norway	Paraguay
99	Nepal	Burundi	Guyana
100	Rwanda	Zimbabwe	Honduras
101	Tanzania	Uganda	Egypt
102	Turkey	Tanzania	Togo
103	Gabon	Rwanda	Sudan
104	Cape Verde	Suriname	El Salvador
105	Suriname	Timor-Leste	Comoros
106	Papua New Guinea	Côte d'Ivoire	Guinea-Bissau
107	Eswatini	Benin	South Sudan
108	Malawi	Mali	Trinidad & Tobago
109	Bahrain	Nicaragua	Argentina
110	Senegal	Botswana	Kyrgyzstan
111	Bangladesh	Bhutan	Mali
112	Egypt	Zambia	Mexico
113	Burkina Faso	Kuwait	Ireland
114	Mozambique	Bahrain	Lesotho
115	Uganda	Kenya	Vietnam
116	Dominican Republic	Luxembourg	Guinea
117	Zambia	Guyana	South Africa
118	Paraguay	Cameroon	Mongolia
119	Angola	Venezuela	Equatorial Guinea
120	Laos	Canada	Bhutan
121	Nigeria	Trinidad & Tobago	Senegal
122	Pakistan	Saudi Arabia	Cameroon
123	Sierra Leone	Oman	Chad
124	Ethiopia	Mozambique	Côte d'Ivoire
125	Lebanon	Lebanon	Suriname
126	Côte d'Ivoire	Namibia	Zimbabwe
127	Bosnia-Herzegovina	Djibouti	Laos
128	Uzbekistan	Pakistan	Timor-Leste
129	Honduras	Mauritania	Cape Verde
130	Niger	Nigeria	Central African Republic
131	Comoros	India	Burundi
132	Algeria	Haiti	Nepal

End of Table C.1

133	Djibouti	Myanmar	Djibouti
134	Madagascar	Niger	Bangladesh
135	Guinea	Algeria	India
136	Zimbabwe	Madagascar	Myanmar
137	Equatorial Guinea	Turkmenistan	Philippines
138	Mali	Burkina Faso	Nicaragua
139	Ukraine	Congo, Rep	Guatemala
140	Tajikistan	Sierra Leone	Uganda
141	Cameroon	Chad	Niger
142	Azerbaijan	Guinea-Bissau	Afghanistan
143	Mauritania	Ethiopia	Burkina Faso
144	Iraq	Papua New Guinea	Zambia
145	Guinea-Bissau	Liberia	Mauritius
146	Sudan	Singapore	Kenya
147	Cambodia	Switzerland	Moldova
148	Nicaragua	Gabon	Tanzania
149	Liberia	Qatar	Rwanda
150	Turkmenistan	Angola	Nigeria
151	Togo	United States of America	Indonesia
152	Congo, Rep	Guinea	Ethiopia
153	Afghanistan	Eritrea	The Gambia
154	Venezuela	Afghanistan	Congo, DR
155	Burundi	Congo, DR	Ghana
156	Chad	Comoros	Papua New Guinea
157	Eritrea	Equatorial Guinea	Sierra Leone
158	Congo, DR	United Arab Emirates	Liberia
159	Central African Republic	South Sudan	Benin
160	Syria	Syria	Mozambique
161	Haiti	Libya	Madagascar
162	Libya	Iraq	Malawi
163	South Sudan	Central African Republic	Haiti
164	Yemen	Yemen	Panama
165	Somalia	Somalia	Costa Rica

Appendix D

Diagnostics and Additional Results

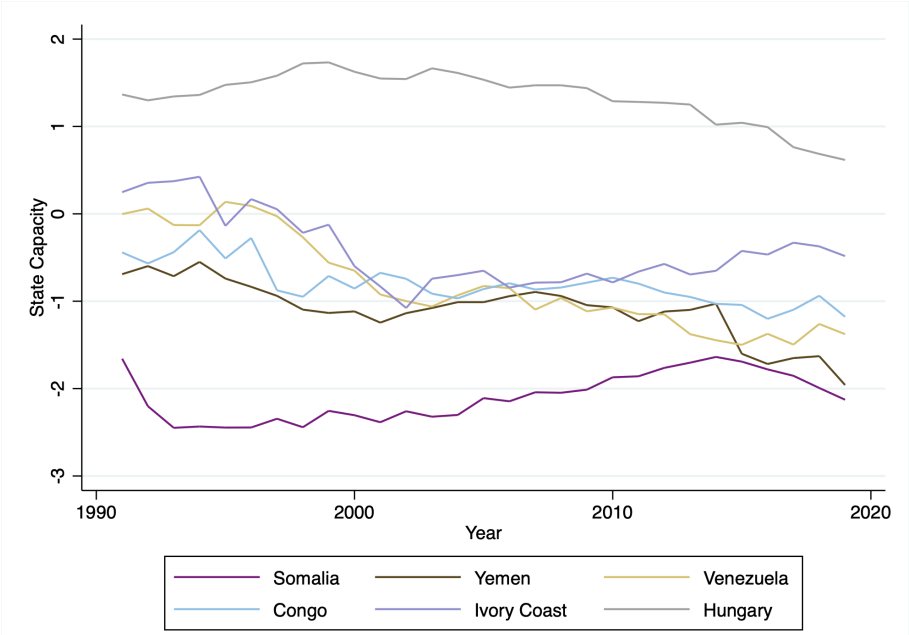
TABLE D.1: Pairwise correlation of State Capacity with base indicators

Indicators	*r*
Bureaucratic Quality	0.9104
Rigorous and Impartial Public Administration	0.8883
Admin. Appointment Decisions Based on Skills	0.8367
Corruption	0.7901
Fiscal Capacity	0.7749
Information Capacity	0.7189
Law and Order	0.7121
Taxes on International Trade as % of Taxes	-0.6815
Census Frequency	0.5825
(log) Tax Revenue as % of GDP	0.5817
(log) Military Expenditures per capita	0.5804
State Authority over Territory	0.5147
Statistical Capacity	0.5116
Taxes on Income as % of Taxes	0.4893
State Antiquity Index	0.4463
(log) Police Officers per 100,000 capita	0.3555
(log) Natural Resource Revenue as % of GDP	
Relative Political Capacity	
(log) Armed Forces Personnel per 1000 capita	
(log) Military Expenditures as % of GDP	

Blanks represent r<.3

FIGURE D.1: Over-time variation in state capacity

(A) Most decreasing countries



(B) Most increasing countries

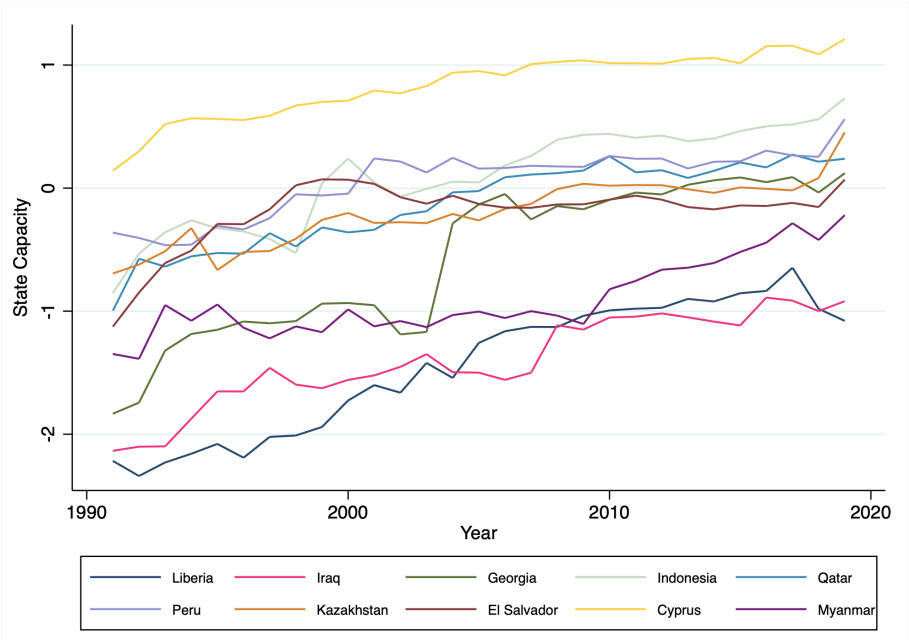


TABLE D.2: Pairwise correlation of other measures with State Capacity's base indicators

Indicators	*State Capacity*	*Government Effectiveness*	*Rule of Law*	*Political Stability*	*Regulatory Quality*	*Fragile States Index*	*State Fragility Index*	*Stateness*	*Government Performance*	*Governance Index*	*Quality of Pub.Adm*
State Capacity	1	0.93	0.91	0.69	0.89	-0.87	-0.79	0.66	0.77	0.72	0.60
icrg_bq	0.9104	0.8801	0.8307	0.6083	0.8018	-0.7885	-0.7374	0.5177	0.5636	0.5101	0.4016
pubadmimpartial	0.8883	0.8370	0.8554	0.6983	0.8281	-0.8049	-0.7341	0.5966	0.7448	0.7130	0.4640
critrecadm	0.8367	0.7781	0.7627	0.5256	0.7467	-0.6983	-0.6192	0.4372	0.6161	0.5943	0.3875
icrg_corruption	0.7901	0.8067	0.8232	0.6278	0.7649	-0.8097	-0.5876	0.5273	0.6142	0.5865	0.4814
fisccap	0.7749	0.7258	0.6995	0.5168	0.7313	-0.6871	-0.6672	0.6052	0.6556	0.6113	0.3703
infcap	0.7189	0.6814	0.6519	0.4799	0.6787	-0.7120	-0.6485	0.5252	0.5652	0.5048	0.3764
icrg_laworder	0.7121	0.7281	0.7726	0.6165	0.6763	-0.7080	-0.6128	0.4235	0.2966		0.4210
tradetax_prctax	-0.6815	-0.6010	-0.5724	-0.4579	-0.6216	0.5978	0.6613	-0.4736	-0.5325	-0.4976	
censusfreq	0.5825	0.5547	0.5251	0.4537	0.5220	-0.5751	-0.5203	0.4654	0.4704	0.4311	
logtaxrev_prcgdp	0.5817	0.5217	0.4971	0.4820	0.5098	-0.5020	-0.5600	0.5817	0.4396	0.4250	
logmilex_capita	0.5804	0.6348	0.6087	0.4072	0.5701	-0.6016	-0.5850				
statautterr	0.5147	0.5058	0.4886	0.6230	0.4586	-0.5329	-0.5626	0.7416	0.4128	0.3769	0.3809
statcap	0.5116	0.4150	0.3595		0.4427	-0.3248	-0.5141	0.5191	0.4972	0.4779	0.5422
incometax_prctax	0.4893	0.3726	0.3549		0.3364						
statehist	0.4463	0.4302	0.3935		0.3870	-0.4133	-0.3314				
logpoloff	0.3555	0.4026	0.3493	0.3431	0.3744	-0.4311	-0.6485	0.3334			
lognatresourcerev_prcgdp											
relpolcap							-0.3074	0.3398			
logmilper_1000capita											
logmilex_prcgdp									-0.3278	-0.3538	

Blanks represent r<.3

TABLE D.3: State Capacity and conflict onset, 1992-2017 (onset1)

	Intrastate conflict onset				
	(1)	(2)	(3)	(4)	(5)
State Capacity[a]	–0.690***	–0.577***	–0.369**	-0.379*	-0.292
	(0.220)	(0.178)	(0.179)	(0.204)	(0.216)
(log) GDP/capita[a]	–0.485	–0.350	–0.789**	–0.788**	–0.795**
	(0.326)	(0.268)	(0.320)	(0.319)	(0.317)
(log) Population[a]	1.279***	1.081***	1.050***	1.052***	1.033***
	(0.227)	(0.166)	(0.157)	(0.153)	(0.161)
Peace Years		–0.290***	–0.268***	–0.267***	–0.270***
		(0.097)	(0.092)	(0.091)	(0.092)
(log) Natural Resource Wealth[a]			0.832***	0.837***	0.809***
			(0.296)	(0.307)	(0.305)
Electoral Democracy[a]				0.050	
				(0.670)	
Liberal Democracy[a]					–0.369
					(0.796)
N	4213	4213	4213	4213	4213

Notes: (a) lagged one year; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Values in parentheses are robust standard errors clustered by country-ID

Coefficients for constants and cubic splines are not shown

TABLE D.4: State Capacity and conflict onset, 1992-2017 (onset5)

	Intrastate conflict onset				
	(1)	(2)	(3)	(4)	(5)
State Capacity[a]	–0.811***	–0.800***	–0.657***	–0.685***	–0.565***
	(0.182)	(0.184)	(0.188)	(0.234)	(0.246)
(log) GDP/capita[a]	–0.006	–0.038	–0.336	–0.334	–0.347
	(0.249)	(0.248)	(0.315)	(0.316)	(0.312)
(log) Population[a]	1.171***	1.161***	1.134***	1.142***	1.111***
	(0.167)	(0.166)	(0.164)	(0.160)	(0.165)
Peace Years		–0.230*	–0.222	–0.221	-0.224
		(0.138)	(0.138)	(0.138)	(0.138)
(log) Natural Resource Wealth[a]			0.586*	0.600*	0.557*
			(0.317)	(0.322)	(0.321)
Electoral Democracy[a]				0.136	
				(0.646)	
Liberal Democracy[a]					–0.420
					(0.790)
N	4213	4213	4213	4213	4213

Notes: (a) lagged one year; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Values in parentheses are robust standard errors clustered by country-ID

Coefficients for constants and cubic splines are not shown

TABLE D.5: State Capacity and conflict onset, 1992-2017 (regional effects)

	Intrastate conflict onset						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State Capacity[a]	–0.816***	–0.821***	–0.838***	–0.527**	–0.533**	–0.451*	–0.462*
	(0.221)	(0.224)	(0.230)	(0.226)	(0.232)	(0.235)	(0.240)
(log) GDP/capita[a]	–0.229	–0.171	–0.243	–0.584	–0.677**	–0.590	–0.689**
	(0.327)	(0.387)	(0.329)	(0.360)	(0.334)	(0.362)	(0.332)
(log) Population[a]	1.295***	1.315***	1.303***	1.185***	1.155***	1.169***	1.139***
	(0.235)	(0.223)	(0.234)	(0.175)	(0.178)	(0.182)	(0.186)
Sub-Saharan Africa[b]		0.116		0.189		0.198	
		(0.308)		(0.246)		(0.244)	
W. Europe & N.America[b]			0.257		0.104		0.170
			(0.612)		(0.604)		(0.629)
Peace Years				–0.108	–0.104	–0.112	–0.108
				(0.099)	(0.100)	(0.099)	(0.100)
(log) Natural Resource Wealth[a]				0.966***	0.964***	0.940***	0.939***
				(0.299)	(0.300)	(0.296)	(0.298)
Electoral Democracy[a]				0.139	0.147		
				(0.732)	(0.739)		
Liberal Democracy[a]						–0.224	–0.222
						(0.860)	(0.898)
N	4213	4213	4213	4213	4213	4213	4213

Notes: (a) lagged one year; (b) dichotomous variable; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Values in parentheses are robust standard errors clustered by country-ID

Coefficients for constants and cubic splines are not shown

TABLE D.6: State Capacity and conflict onset, 1992-2013 (ethnic fractionalization)

	Intrastate conflict onset				
	(1)	(2)	(3)	(4)	(5)
State Capacity[a]	–0.956***	-0.624***	–0.541**	–0.644***	–0.557**
	(0.196)	(0.237)	(0.244)	(0.237)	(0.239)
(log) GDP/capita[a]	0.019	–0.289	–0.289	–0.190	–0.189
	(0.291)	(0.325)	(0.321)	(0.314)	(0.312)
(log) Population[a]	0.943***	0.828***	0.801***	0.845***	0.819***
	(0.191)	(0.168)	(0.171)	(0.160)	(0.162)
Peace Years		–0.105	–0.105	–0.093	–0.092
		(0.139)	(0.139)	(0.141)	(0.140)
(log) Natural Resource Wealth[a]		0.579*	0.560*	0.594**	0.571**
		(0.296)	(0.290)	(0.271)	(0.266)
Electoral Democracy[a]		–0.327		–0.357	
		(0.686)		(0.657)	
Liberal Democracy[a]			–0.752		–0.813
			(0.849)		(0.813)
Ethnic Fractionalization[a]				0.778**	0.790**
				(0.390)	(0.384)
N	3327	3327	3327	3327	3327

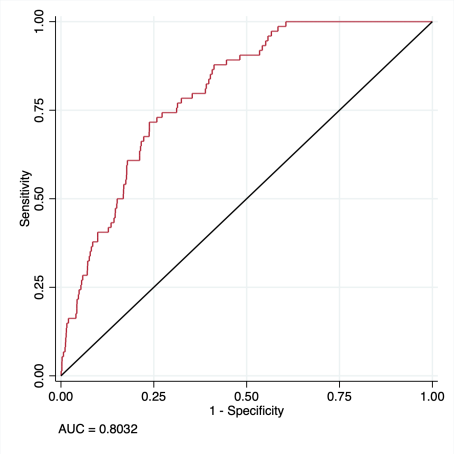
Notes: (a) lagged one year; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Values in parentheses are robust standard errors clustered by country-ID

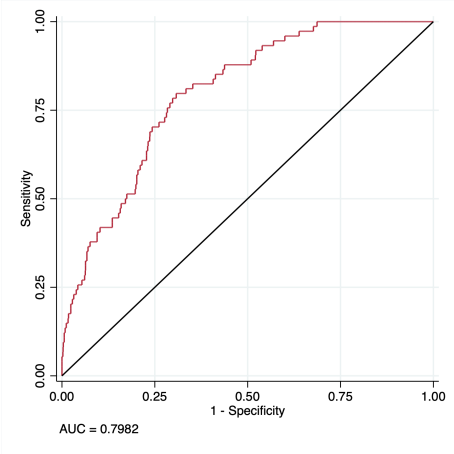
Coefficients for constants and cubic splines are not shown

FIGURE D.2: Receiver operating characteristic (ROC) curves

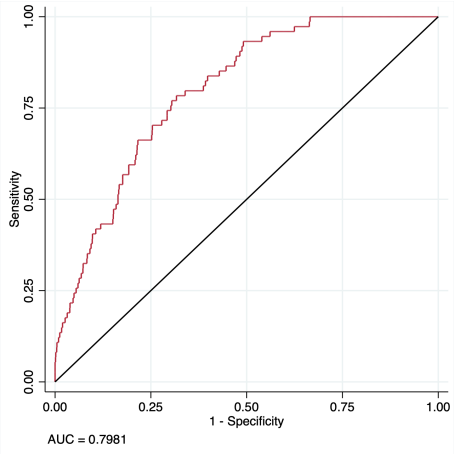
(A) State Capacity



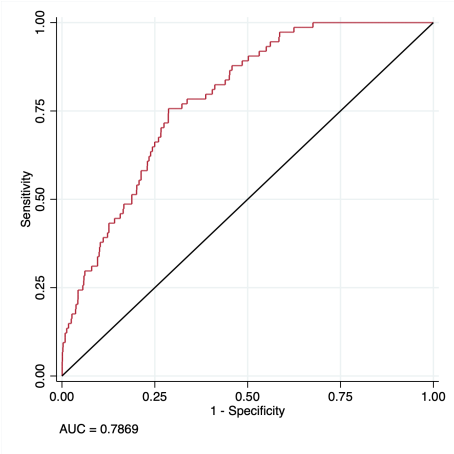
(B) Government Effectiveness



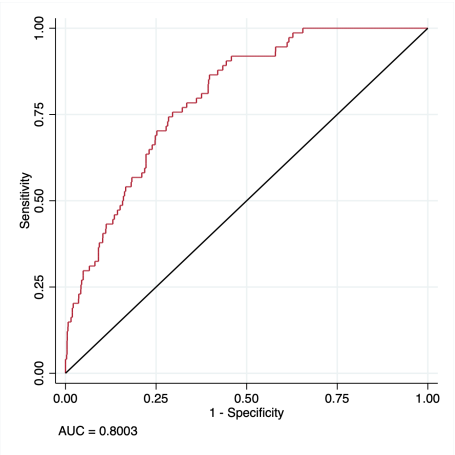
(C) Regulatory Quality



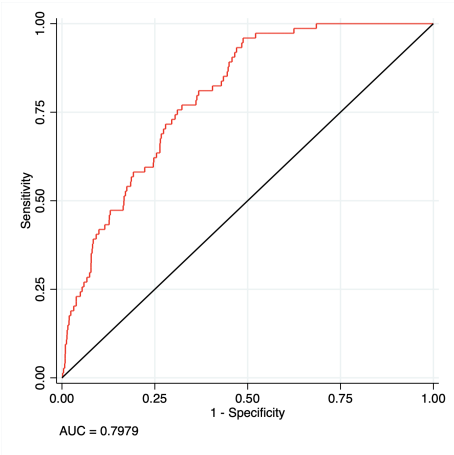
(D) Rule of Law



(E) State Fragility Index



(F) Fragile States Index



Appendix E

Rscript and Do-file

The statistical computing in this thesis was done using both R and Stata. For multiple imputation and for generating most figures, R was applied. When conducting the factor analysis and regression models, Stata was the preferred software. All syntax and files can be provided upon request: charltan@student.sv.uio.no.