



Calhoun: The NPS Institutional Archive
DSpace Repository

NPS Scholarship

Publications

2019-06

Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research

Hanson, Jonathan K.; Sigman, Rachel

Hanson, Jonathan K., and Rachel Sigman. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." (2019).
<https://hdl.handle.net/10945/64717>

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research

Jonathan K. Hanson
Ford School of Public Policy
University of Michigan

Rachel Sigman
Naval Postgraduate School

June 2019

Abstract

State capacity is a core concept in political science research, and it is widely recognized that state institutions exert considerable influence on outcomes such as economic development, civil conflict, democratic consolidation, and international security. Yet, researchers across these fields of inquiry face common problems involved in conceptualizing and measuring state capacity. This paper examines these conceptual issues, identifies three common dimensions of state capacity, and uses Bayesian latent variable analysis to assess the extent to which these dimensions are discernible in available indicators of state capacity. The analysis results in a new general-purpose measure of state capacity with demonstrated validity for use in a wide range of empirical inquiries. The paper also establishes the difficulty of achieving empirical distinctions between core dimensions of state capacity. It is hoped that this project will provide effective guidance and tools to researchers studying the causes and consequences of state capacity.

In the influential volume, *Bringing the State Back In*, Evans, Rueschemeyer, and Skocpol (1985) noted a surge of interest in the state as an actor. This interest has not abated in the years since. It is widely recognized that state institutions exert considerable influence on outcomes including economic growth, human development, civil conflict, international security, and the consolidation of democracy. Along with the proliferation of theories containing state capacity as an explanatory variable, however, has come divergence in how it is conceptualized, impeding our ability to compare findings and expand our understanding of its roles.¹

The difficulty of measuring state capacity empirically, however conceptualized, magnifies this problem. We seek to address three common challenges that researchers face in selecting and employing quantitative measures of state capacity. First, absent clear definition of the concepts underlying state capacity, researchers may select measures that do not effectively represent the dimension(s) of state capacity most relevant to their research (Soifer, 2008). Second, geographic and temporal coverage is often sparse. Third, it is difficult for researchers to employ measures that are distinct from concepts of interest such as economic development or regime type.

In this paper, we argue for a narrow definition of state capacity to clarify its distinctiveness from similar-sounding concepts. This discussion facilitates the identification of three core dimensions of state capacity: extractive capacity, coercive capacity, and administrative capacity. In the third section, we discuss the challenges facing researchers in developing and employing empirical measures of state capacity. To address these challenges, we use a Bayesian latent variable analysis that employs 22 indicators of state capacity. The resulting Capacity measure provides annual estimates of the levels of state capacity for all countries included in the Polity dataset (Marshall and Jaggers, 2009) from 1960-2015. We then assess the new measure with empirical tests and discuss the challenges of empirically disaggregating state capacity for quantitative analysis.

¹Following Skocpol (1985), we define the state as a set of coercive, extractive, legal and administrative organizations responsible for providing order and other public goods to a society.

1 Defining State Capacity

The meaning of state capacity varies considerably across the literature in political science. Further complications arise from an abundance of terms that refer to closely-related attributes of states: state strength or power, state fragility or failure, infrastructural power, institutional capacity, political capacity, quality of government or governance, and the rule of law. In practice, even when there is clear distinction between these terms at the conceptual level, data limitations frequently lead researchers to use the same empirical measures for differing concepts.

To prevent conceptual creep and provide a firm basis for an empirical approach that will produce useful results for a wide range of research, we argue for a minimalist definition of state capacity: the ability of state institutions to effectively implement official goals (Sikkink, 1991; Fukuyama, 2013). This definition avoids normative conceptions about what the state ought to do or how it ought to do it. Instead, we adhere to the notion that capable states may regulate economic and social life in different ways and may achieve these goals through various types of relationships with social groups.²

Three principles guide us in operationalizing this definition for the purpose of creating a quantitative measure of state capacity suitable for cross-national research. First, we focus only on core functions of the state rather than on the whole spectrum of potential government action. Second, we steer away from entanglement with related concepts of interest in political science research. Third, while we acknowledge the utility of measuring territorial or sectoral variation in a state's capacity, we seek to develop a national-level measure of state capacity suitable for cross-national comparative research.

Although we use a narrow definition of state capacity, we recognize that the concept is multi-dimensional. As Levi notes, “good analysis requires differentiating among the features of the state in order to assess their relative importance; the state becomes less than the sum of its parts”

²For recent discussions and debates about the definition of state capacity, see Centeno et al. (2017) and Lindvall and Teorell (2016). Both works suggest that state capacity ought to be construed more narrowly as something that the state possesses, but does not necessarily use. Though they use different terms (Centeno et al. (2017) suggest “state performance” and Lindvall and Teorell (2016) suggest “state power”), both definitions relate to the extent to which the state fulfills its functions and accomplishes goals. We maintain use of the term state capacity for consistency with the bulk of the literature.

(2002: 34). Yet, despite the widespread recognition of the state as a multi-dimensional concept, there is considerable divergence across the literature on how to disaggregate state capacity at the conceptual level.³ We focus on those core functions most commonly deemed necessary for modern states. These functions include protection from external threats (Tilly, 1990), the maintenance of internal order (Huntington, 1968), the administration and provision of public goods necessary to sustain economic activity (North, 1981, 1990) and the extraction of revenue (Tilly, 1990; Levi, 1988).

Focusing on these core functions not only steers clear of normative questions about what states should do but also avoids overlap with concepts such as “good governance,” institutional quality, or state autonomy that may be theoretically or empirically related to state capacity as defined in this article. Although institutional quality, meaning the extent to which governments are free of corruption, impartial and/or accountable to citizens, may relate closely to the state’s ability to fulfill its core functions (Dahlström and Lapuente, 2017; Taylor, 2011; Bäck and Hadenius, 2008), the presence of corruption or partiality in the state does not necessarily preclude success in achieving official goals. Moreover, as Fukuyama (2013) notes, measures of state capacity that incorporate such concepts impede our ability to examine the relationship between, for example, regime type and capacity. For example, Ang (2016) illuminates how ‘low-quality’ institutions in China supported the state’s ability to successfully pursue reforms that helped to boost revenue, expand public goods services and support economic growth.

Finally, although state capacity varies across territory (Giraudy and Luna; Luna and Soifer, 2017; Harbers, 2015; Soifer, 2012) and policy sectors (Whitfield et al., 2015; Krasner, 1978), national-level measures remain important for both theoretical and practical reasons. Approaches that seek to measure state capacity by working up from the policy sector or sub-national level, moreover, face a number of thorny problems. First, as Skocpol (1985) remind us, states must exist in, and are recognized by, an international system. A subnational or sectoral focus risks oversight of

³For example, Brautigam (1996) divides state capacity into four dimensions: regulatory, administrative, technical and extractive capacities. Besley and Persson (2009) identifies two functional categories: legal and fiscal capacity. For Bäck and Hadenius (2008), state capacity consists of stateness and administrative capacity. Sobek’s (2010) dimensions of state capacity include economic development, extractive capabilities and bureaucratic quality. Centeno et al. (2017) divide state performance into order and reach, economic development and inclusion and equity. There are many more examples. In their approach to state power, Lindvall and Teorell (2016) distinguish between human, material and informational capacities of the state.

the important ways that the international system shapes states' capacities (Clapham, 1996; Jackson and Rosberg, 1982; Herbst, 2000) and determines the nature of external threats (Tilly, 1990).

Second, measures of sub-national or sector-specific capacity are more likely to conflate the results of political *choices* about policy priorities or resource allocation as variation in state capacity. A large body of work links distributive decisions to political priorities (Golden and Min, 2013) and shows that logics of distribution vary considerably across countries (Kramon and Posner, 2013). Wealthier neighborhoods may get faster police response times, areas with politically important constituencies may receive more health clinics, and so forth. We cannot assume that unevenness in delivery of public services is an indicator of unevenness in state capacity.

Third, from an empirical standpoint, disaggregation aimed at measuring unevenness of state capacity across different areas ultimately leads to reductionism, since policy sectors or localities may be divided into ever smaller components. Related to this problem, measuring numerous features of the state across countries and over time increases data collection burden dramatically and threatens comparability if state scope varies. Even if one were to collect data on a wide range of policy areas or subnational regions across countries and over time, it is not clear how to aggregate these data into a broad measure of state capacity of the kind that is most useful to many researchers.

2 Dimensions of State Capacity

Following the principles laid out in the previous section we seek to identify dimensions that: 1) focus on the core functions of the central state; and 2) avoid conflation with other concepts of interest. We thus concentrate on three dimensions of state capacity that are minimally necessary to carry out the functions of contemporary states: extractive capacity, coercive capacity, and administrative capacity. These three dimensions, described in more detail below, accord with what Skocpol identifies as providing the “general underpinnings of state capacities” (1985: 16): plentiful resources, administrative-military control of a territory, and loyal and skilled officials. Our objective of developing broad measures of state capacity for comparative political research is best served by concentrating on these underpinnings.

Perhaps nothing is more central to the concept of state capacity than raising revenue. North defines the boundaries of the state in terms of its ability to tax constituents (1981: 21), while Levi (1988) and Tilly (1990) make a direct connection between a state's revenue and the possibility to extend its rule. Raising revenue is not only a critical function of the state, but it also encompasses a particular set of capacities that are foundational to state power. In particular, states must have the wherewithal to reach their populations, collect and manage information, possess trustworthy agents to manage the revenue, and ensure popular compliance with tax policy. Empirically, taxation is associated with property rights (Besley and Persson, 2009), the reach of the state (Harbers, 2015) and state "legibility" (Lee and Zhang, 2017). We characterize this particular group of capacities as extractive capacity.

Like extractive capacity, coercive capacity is also central to the definition of the state, particularly in the Weberian tradition that defines the state as the organization possessing a monopoly on the legitimate use of force within its territory (Weber, 1918). Coercion relates directly to the state's ability to preserve its borders, protect against external threats, maintain internal order, and enforce policy. To achieve broader policy goals, including the collection of revenue, a state must be able to tame violence (Bates, 2001) by possessing the force necessary to contain threats throughout its territory, or at least convince its rivals that this is the case. While coercion is not the only way to maintain order and evoke compliance from the population (Levi, 1988), it represents a key aspect of the ability of states to survive and implement policies.

Administrative capacity is a broader dimension that includes the ability to develop policy, the ability to produce and deliver public goods and services, and the ability to regulate commercial activity. Effective policy administration requires technical competence, trusted and professional state agents, monitoring and coordination mechanisms, and effective reach across the state's territory and social groupings. In particular, Weber (1978) emphasizes the importance of autonomous and professional bureaucracies that legitimize the authority of the state, manage complex affairs, and ensure efficiency, including the control of corruption.

Although we believe these three dimensions of state capacity represent analytically distinct features of states, there are a number of studies that theorize the interrelationships between these dimensions. Tilly (1990), for example, explains how leaders extracted revenue to develop coercive

capacity to protect and expand control of territory. In a recent analysis of state capacity indicators, Rogers (2014) argues in favor of this intuition by suggesting that income tax serves as a valid overall indicator of state capacity. Another connection between capacity dimensions is found in literature on rentier states, where a decreased need to extract resources from taxpayers is associated with an erosion of the state’s capacity to administer public policy (Chaudhry, 1997; Hertog, 2011). There is not always a clear unidirectional relationship between these dimensions. Fortin-Rittenberger (2014), for example, finds that administrative and coercive capacities can be both positively and negatively related, depending on which aspect of coercive power is considered. Finally, recent attention to “informational capacity” (Brambor et al., 2016; D’Arcy and Nistotskaya, 2017) and legibility (Lee and Zhang, 2017) suggests the possibility that certain underlying capacities may make states more capable across the extractive, administrative and coercive spheres. That these types of interrelationships are increasingly prominent in studies of state capacity suggests the potential for difficulty in empirically distinguishing between the three dimensions despite their conceptual distinctiveness. This is an issue we return to below.

3 Measurement Strategies and Challenges

Researchers use a growing variety of indicators to measure state capacity, but it is not always clear if indicators are relevant to the dimension(s) of primary theoretical interest, or if they adequately consider the range of state capacity dimensions that may relate to their variable of interest.⁴ In this section, we examine and assess indicators used to measure the three dimensions presented above. In addition to examining the conceptual validity of each indicator, we also assess the extent to which indicators overlap with other concepts of interest as well as their temporal and geographic coverage. Where possible we look to combine both “objective” data reflecting actual practices such as the collection of taxes and the frequency of censuses as well as more subjective measures based typically on expert perceptions.

⁴Saylor (2013), in particular, points to the need to consider all dimensions of a concept being measured.

3.1 Measures of Extractive Capacity

Given the fundamental nature of revenue extraction to the state, many researchers utilize data on government revenue collections as a measure of state capacity.⁵ Tax revenue data are available for most countries from the early-1970s onwards, generally from the IMF's *Government Finance Statistics*.⁶ Data on different types of revenues are usually expressed as a raw amount, as a proportion of GDP, or as a proportion of total revenue collected. As Lieberman (2002) explains, there are many factors to consider when selecting revenue indicators that are appropriate for a particular purpose.⁷

Aggregate revenue, for example, is a noisy indicator of extractive capacity. For many states with relatively high extractive capacity, the level of tax revenue collection reflects a policy choice rather than the state's extractive capacity. Additionally, different types of revenue vary significantly in terms of their administrative complexity and their political implications. As Lieberman (2002) and Rogers (2014) argue, the revenue sources that are most likely to capture concepts related to state capacity include income, property and domestic consumption taxes. These taxes are more administratively complex, requiring higher levels of record-keeping, transparency, and a more sophisticated bureaucratic apparatus than other revenue sources. Taxes on international trade, on the other hand, are much easier to collect and, like rents from mineral resources, do not require significant enforcement capacity (2002: 98).⁸

Our strategy with respect to revenue measures is twofold. First, we use tax revenues as a proportion of GDP to capture overall extractive capacity. We exclude non-tax revenues for the reasons Lieberman identifies. Second, we expect that the mixture of tax revenues – specifically taxes on income and taxes on trade – provides information about *both* the extractive and administrative capacities of the state. Given a particular level of taxation, the greater the proportion of tax revenue that comes from income taxes, the higher the expected level of administrative capacity.

⁵See, for example, Besley and Persson (2009) and Dincecco (2017).

⁶Prichard et al. (2014) have usefully standardized and compiled tax data from IMF country records. Tax data are also available from other sources such as the OECD. Lieberman (2002) provides an overview of taxation data as a measure of state capacity.

⁷See also Rogers (2014).

⁸Chaudhry (1997) underscores this point, showing that the bureaucracy's administrative capacity in Saudi Arabia actually declined as the collection of oil rents and trade taxes increased with their oil production.

The opposite should be true with respect to the proportion of revenue that comes from taxes on trade, which are administratively easy to collect. We thus use the proportion of tax revenues –as opposed to taxes as a proportion of GDP– that come from these two sources as measures of the administrative capability of the state’s extractive efforts.

Other indicators tap into the extractive capacity dimension. The World Bank’s (2011a) Country Policy and Institutional Assessment (CPIA) index, for example, includes a rating of the Efficiency of Revenue Mobilization for 74 countries starting in 2005. Finally, we expect that some of the indicators that are logically related to the dimensions of coercive and administrative capacity will also provide information about extractive capacity. For example, a state’s ability to collect information about its citizens is relevant for extractive capacity, and it may be measured in part by the frequency with which a country has a census (Centeno, 2002b; Soifer, 2013; Hanson, 2015). Similarly, the World Bank’s (2012) Bulletin Board on Statistical Capacity measures the quality of statistical systems in developing countries in 1999 and from 2004-present.

3.2 Measures of Coercive Capacity

Researchers seeking to measure coercive capacity may turn attention to military size or sophistication, as well as attributes of the state thought to promote the maintenance of order. Data on military expenditures, military personnel, and security forces are available from large-N datasets such as the World Development Indicators, the Stockholm International Peace Research Institute, and the Correlates of War (Singer et al., 1972). Coverage and reliability for these measures is generally quite good for most countries in the period 1960 to the present. The relationship between coercive force and a state’s coercive capacity, however, is not necessarily straightforward (Hendrix, 2010; Kocher, 2010; Soifer and vom Hau, 2008). States that have the capacity to maintain order might have effective military and/or security forces, although there are countries that maintain order with little or no military. A large military force, moreover, may be a sign of war or insecurity, both of which could deplete state capacity. In this project, we use the log value of military expenditures per million in population and the number of military personnel per thousand in the population (Singer et al., 1972; World Bank Group, 2011b) as indicators of military capacity. We supplement these indicators with estimates of forces intended for the maintenance of internal order

including the size of paramilitary forces compiled from Sudduth (2016), extended by the authors from the annual listings in *Military Balance*. We also include a measure of the size of the police force obtained from the United Nations Office on Drugs and Crime.

We also include two expert-coded indicators of coercive capacity. From the Bertelsmann Transformation Index (Bertelsmann Stiftung, 2006), we adopt a measure that assesses the degree to which the state has a monopoly on the use of force. We also include ratings from The Political Risk Services' (PRS) International Country Risk Guide on "law and order" which assesses the strength and impartiality of the legal system, and the popular observance of the law (Howell, 2011).

Finally, we use a set of indicators that tap the dimension of coercive capacity by capturing the state's overall level of political institutionalization or presence in the territory. The first of these is *Anocracy*, a measure calculated from the Polity dataset (Marshall and Jaggers, 2009) based on insights from Gates et al. (2006) and Vreeland (2008). Anocracy is an index that ranges from 0-12, and it increases in the extent to which states show consistency in their political institutions.⁹ Second, we use V-Dem's (Coppedge et al., 2019; Pemstein et al., 2019) measure of State Authority over Territory, which measures the percent of territory controlled by the central state. Third, we use the State Antiquity Index developed by Bockstette et al. (2002) to measure the state's age.¹⁰ According to this measurement strategy, states with a longer history of government institutions and control of their territories are assigned higher scores.

3.3 Measures of Administrative Capacity

Since administrative capacity is a broad dimension of state capacity, a number of different measurement strategies exist. A common way to measure administrative capacity is to look at the outcomes of public goods and service delivery such as the percentage of children enrolled in primary schools, infant mortality rates, or literacy rates. These measures are attractive for their broad coverage and comparability, but assessing capacity based on measures of this kind poses several problems.

⁹Specifically, states which consistently show the features of either institutionalized democracy or autocracy score higher, while those that contain a mixture of features or lower levels of score lower.

¹⁰We extend the measure developed by Bockstette et al. (2002) with an annual coding from 1959 through 2015. We then combine this annualized measure with the original measure's "stock" of the state to produce annualized scores for 1960-2015.

First, as discussed above, a state may not prioritize the particular outcome being measured, such as schooling or health or infrastructure. Second, using these measures may compromise analytical leverage, since these types of outcomes are attributable to a number of different factors such as levels of economic development, the nature of the political regime, or participation in international programs with policy conditions.

Among indicators of administrative capacity, two of the most popular are the Government Effectiveness rating from the Worldwide Governance Indicators (Kaufmann et al., 2003) and the International Country Risk Guide’s (ICRG) Bureaucratic Quality rating (Political Risk Services, 2010). Both measures have come under scrutiny. The WGI, for example, are frequently criticized for their aggregation procedures and for the fuzzy analytical boundaries that characterize their different governance indices.¹¹ The ICRG Bureaucratic Quality ratings, on the other hand, may be prone to measurement errors based on analyst perceptions of economic or social outcomes rather than bureaucratic quality *per se* (Rauch and Evans, 2000; Henisz, 2000). In our case, using the WGI scores would be especially problematic because the set of constituent indicators overlaps with ours. We do, however, include the ICRG Bureaucratic Quality rating in our analysis.

We also include several indexes related to administrative capacity from various sources: Administrative Efficiency (Adelman and Morris, 1967), the Weberianness index (Rauch and Evans, 2000), From the World Bank’s CPIA index there are two relevant ratings: Quality of Budgetary and Financial Management and Quality of Public Administration. None of these ratings covers a long period of time, but the combination covers significant portions of the 1960-2015 time period with at least one indicator.¹² Similar to the Weberianness measure is a measure of impartial public administration developed by V-Dem (Coppedge et al., 2019; Pemstein et al., 2019).

Additionally, we derive a measure of census frequency calculated with data on country censuses provided by the International Programs Center of the U.S. Census Bureau.¹³ As argued in Centeno

¹¹There has been quite a bit of debate about the validity, reliability, and aggregation of the WGI. For an overview and response to critiques see Kaufman et al. (2007).

¹²All except for Administrative Efficiency and the Weberianness index cover countries only in the 2000s. We code Administrative Efficiency as covering the years 1960-1962 and Weberianness as covering the period 1970-1990 based on the scholarly objectives of their creators.

¹³We have annualized this measure, which ranges from 0.15 to 2.39, 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.

(2002b) and Soifer (2013), countries that can conduct censuses have not only the capacity to collect information exhibit higher levels of territorial reach. For additional indicators of informational capacity, we use the measure of informational capacity developed by Brambor et al. (2016) covering 166 countries and the World Bank’s Statistical Capacity measure.

3.4 Indicators overall

Our goal was to assemble a group of indicators best suited to represent, in varying ways, the three theorized dimensions; allow for broad coverage, and, to the extent possible, avoid measures that are likely to capture other concepts of interest. Altogether, we employ 22 different indicators related to the three key dimensions of state capacity. The data span 56 years (1960-2015) and up to 163 countries in a given year.¹⁴ The indicators employed in this analysis are listed in Table 1. By adopting a latent variable analysis of the kind employed to assess measures of democracy (Treier and Jackman, 2008) and governance (Arel-Bundock and Mebane, 2011; Bersch and Botero, 2011) we can use these multiple measurements of the same underlying concepts, even if noisy, to gain information about the distribution of the latent parameters that generate the observed indicators.

4 Latent Variable Analysis

We employ the latent variables estimation approach developed by Arel-Bundock and Mebane (2011), hereinafter ABM, that uses Bayesian Markov-Chain Monte Carlo (MCMC) techniques to identify underlying factors.¹⁵ This technique, based on earlier work by Lee (2007), has many advantages over traditional factor analysis, including robustness to missing data. By incorporating indicators of state capacity drawn from multiple sources, we seek to provide annual measures of state capacity for the set of all countries that appear in the Polity dataset (Marshall and Jaggers, 2009) during the 1960-2008 time period.

Specifically, using the notation of ABM with minor changes, each observed indicator x_k for

¹⁴The number of countries is different in each year based on the sample definition using Polity IV data on the existence of countries. In some cases data for additional countries or years are available but not included because of the way we have defined the sample.

¹⁵We are very grateful to Arel-Bundock and Mebane for making their programming code for this method available.

Table 1: Indicators of State Capacity

Variable	Countries	Years
Administrative Efficiency (Adelman and Morris, 1967)	69	1960-1962
Anocracy (calculated from Polity IV)	175	1960-2015
Bureaucratic Quality (Political Risk Services)	148	1982-2015
Census Frequency (calculated from UN 2011)	179	1960-2015
Efficiency of Revenue Mobilization (World Bank CPIA)	74	2005-2015
Information Capacity (Brambor et al., 2016)	166	1960-2015
Law and Order (Political Risk Services)	174	1984-2015
Military Personnel per 1,000 in population (COW)	171	1960-2015
(Log) Military Spending per million in population (COW)	168	1960-2015
Monopoly on Use of Force (Bertlesmann Transformation Index)	127	2003-2015
(Log) Paramilitary Personnel per 1000 in population	164	1961-2015
Police Officers per 1000 in population (UN)	122	1973-2015
Quality of Budgetary and Financial Management (World Bank CPIA)	74	2005-2015
Quality of Public Administration (World Bank CPIA)	74	2005-2015
Rigorous and Impartial Public Administration (V-Dem v9)	177	1960-2015
State Antiquity Index, based on Bockstette et al. (2002)	162	1960-2015
State Authority over Territory (V-Dem v9)	174	1960-2015
Statistical Capacity (World Bank)	134	2004-2015
Taxes on Income as % of Revenue (IMF, WDI)	152	1970-2015
Taxes on International Trade as % Revenue (IMF, WDI)	155	1970-2015
Total Tax Revenue as % GDP (IMF, WDI, OECD)	152	1960-2015
Weberianess (Rauch and Evans, 2000)	34	1970-1990

country i in time t is a linear function of J latent variables and a disturbance ϵ_k :

$$x_{kit} = c_k + \sum_{j=1}^J \lambda_{kj} \xi_{jit} + \epsilon_{ki} \quad (1)$$

In Equation 1, ξ_{jit} is the latent value of the j th dimension of state capacity for country i in time t , and λ_{kj} is the linear effect of the j th dimension on the observed indicator x_k . Overall, then, the various observed indicators are linear functions of the latent values of state capacity in each dimension measured with some error. Since there are k observed indicators measured in many countries over several years, we have multiple data points with which to obtain the posterior distributions of the latent parameters.¹⁶ We assign standard normal priors to the latent factors. The intercepts c_k have independent, diffuse normal priors, and the disturbance terms ϵ_k have independent uniform priors with mean zero. In general, diffuse normal priors were applied for each λ_{kj}

To facilitate identification, one of the parameters λ_{kj} is fixed at 1 for each of the J dimensions in the analysis. In these cases, the intercepts c_k are fixed at 0. Additionally, truncated (positive) normal priors were applied to facilitate identification where we had a strong prior belief that the relationship between a given indicator (x_k) and the parameter representing Capacity (ξ_j) is positive. For example, in our main model truncated, normal priors are applied in the following cases: census frequency, state antiquity, taxes on income, Weberianness, the V-Dem public administration measure, PRS law and order, and the administrative efficiency rating of Adelman and Morris (1967).¹⁷

The MCMC is implemented in JAGS (Plummer, 2010) through the package rjags (Plummer, 2012) for R statistical software. The algorithm tours the parameter space specified by the sets of equations represented by Equation 1. Successive draws lead to descriptions of the posterior distributions of the remaining parameters that produce the observed indicators of state capacity. A typical MCMC run included five chains with an adaptation phase of 5,000, a burn-in phase of

¹⁶The greater the number of observed indicators, the more information we have about the values of latent dimensions of state capacity in country i at time t . The larger the number of country-years, the more information we have to uncover λ_{kj} , the effect of dimension j on indicator k , which is treated as constant over time.

¹⁷Otherwise, some chains would simply take on the opposite signs of other chains.

10,000 iterations, and a sampling phase of 5,000 iterations. Samples were thinned with a setting of 5 to alleviate memory/storage constraints.

Since we do not know whether the three theorized dimensions are discernible in the data, we run multiple analyses, letting the number of dimensions J range from 1 to 3. The parameter estimates that emerge from choosing a particular number of dimensions, furthermore, need not bear any particular relationship to the theoretical dimensions we describe. As with traditional factor analysis, we rely on analysis of which indicators align with the resulting parameters to interpret the dimensions. One possibility is that each successive dimension captures more marginal aspects of variation in the observed indicators rather than clear dimensions.

5 Latent Variable Analysis Results

In repeated tests, we found that a one-dimensional model ($J = 1$) was the only model to converge consistently. Attempts to identify a second or third dimension did not bear fruit. Typically, the different chains would fail to converge, and the posterior distributions for some parameters would exhibit strong non-normality. Consequently, we present results that reflect a single, latent dimension that we call Capacity.

We believe that there is a very logical interpretation for this outcome: the extractive, coercive and administrative dimensions of state capacity, though distinct conceptually, are interrelated in practice. As discussed above, extractive capacity supports both coercive power and provides the resources needed to sustain a sophisticated administrative bureaucracy. Likewise, states that lack coercive and administrative capabilities are likely to find revenue extraction more difficult. Finally, although state coercion can take many different forms, some of them very simple, coercive power is facilitated by a well-organized, administratively-sophisticated coercive apparatus. These interrelationships make it difficult at the empirical level to disaggregate state capacity into separate dimensions.

5.1 Exploring the New Measure

With 8,254 observations in total, the Capacity measure is much broader in scope than the most commonly used general measures of state capacity. For example, our data exceed the International Country Risk Guide’s measure of Bureaucratic Quality in coverage by over 2,000 observations and with 20 additional years, and they expand upon the estimates from Hendrix (2010) by over 5,000 observations and 24 years. A fuller list appears in the supplemental material. The mean of Capacity is .25 and the standard deviation is .95.

To understand what factors are driving the Capacity estimates, we first examine their correlation with the observed indicators included in the estimation procedure. This analysis, which is presented in Table 2, helps determine how well the state capacity measures align with the three core dimensions described theoretically.

Table 2: Correlation of Capacity with Base Indicators

Indicator	r	N
Statistical Capacity	0.82	1492
Bureaucratic Quality	0.80	4089
Law and Order	0.80	4089
(Log) Military Spending per million in population	0.78	7595
Monopoly on Use of Force	0.77	1247
Rigorous and Impartial Public Administration	0.76	8061
Quality of Public Administration	0.75	724
Quality of Budgetary and Financial Management	0.71	724
State Authority over Territory	0.69	8046
Administrative Efficiency	0.68	199
Efficiency of Revenue Mobilization	0.67	724
Information Capacity	0.65	3440
Total Tax Revenue as % of GDP	0.65	6413
Weberianness	0.59	714
Census Frequency	0.57	8201
Anocracy	0.51	8254
Taxes on Income as % of Revenue	0.50	5511
State Antiquity Index	0.43	8032
Military Personnel per 1,000 in population	0.36	7795
(Log) Paramilitary Personnel per 1,000 in population	0.12	5384
Police Officers per 1,000 in population	0.07	1569
Taxes on International Trade as % of Revenue	-0.68	5563

Capacity appears to be a general-purpose measure of state capacity that draws from indicators

representing all three theorized dimensions. The indicators most strongly associated with Capacity are: the World Bank's measure of Statistical Capacity ($r = .82$); the PRS Bureaucratic Quality ($r = .80$) and Law and Order ratings ($r = .80$); military spending ($r = .78$); the BTI monopoly on use of force measure ($r = .77$); and the V-Dem measures of public administration ($r = .75$ and $.71$). Most of the indicators are correlated with Capacity at the .5 level or greater (or less than -.5 in the case of taxes on trade).

The indicators with weakest correlation to Capacity are the measures related to military, paramilitary, and police personnel. Since other measures of coercive capacity are strongly correlated with Capacity, the pattern appears to be limited to security personnel. A few explanations seem plausible. First, the Capacity measure misses aspects of coercive capacity that arise from state employment of security personnel. Second, rulers of weaker states, or those engaged in conflict, tend to expand their security forces in response to this weakness, thereby further weakening the relationship. Third, it is not the numbers of security personnel that matter but their level capability as measured by their administrative organization.

5.2 Validity Checks

The broader coverage of countries and years is welcome, provided that the measures perform well. The goal of this section is to investigate whether the measures behave in the expected manner, and whether they will be useful for investigating theoretical questions regarding state capacity. Following guidance from Adcock and Collier (2001), Gerring (2011), Seawright and Collier (2014) and McMann et al. (2016), we examine the new measure in terms of its face validity, content validity, convergent validity and nomological validity.

Figure 1 displays the mean and standard deviation of each country's Capacity posterior distribution in the year 2015, ranked from the highest to the lowest on the measure. In terms of face validity, the countries we might expect to have strong state capacity are found to have higher scores, while those that are experiencing or have recently experienced war or have notoriously weak capacity are found to have the lower scores. That Singapore ranks among the fifteen highest Capacity scores help us to know that these measures do not capture concepts more closely related

to democratic governance than to capacity itself. At the lower end of the scale, we see states such as Somalia, Yemen and Central African Republic that are embroiled in conflict, lacking state structures, or both. Plots similar to Figure 1 for a broader range of years are included in the supplemental materials.

Figure 2 plots Capacity scores for all countries in the dataset, with 1960 scores on the x-axis and 2015 scores on the y-axis. As theory would predict, the relationship between the Capacity variables in different years is strongly positive. Most countries starting with high scores in 1960 also have high scores in 2015. The countries deviating from this pattern permit us to examine specific cases to assess validity of changes over time. As can be seen in the figure, the countries where Capacity rose the most were Singapore, Chile, Oman, Uruguay, Ethiopia and Cyprus. Countries where Capacity decreased the most include Venezuela, Iraq, Syria, Libya, Somalia and Central African Republic.

Chile's experience, as described by Garretón and Cáceres (2003), began with Pinochet's expansion of the coercive apparatus, followed in the democratic period beginning in 1990 by a series of administrative reforms designed to modernize administrative institutions and management structures. In Singapore, thanks to support from a broad coalition of social groups (Crone, 1988), the dominant People's Action Party has grown and maintained its strong capacity to regulate social and economic life not only through coercive means, but also through its skilled and efficient bureaucracy. Steady growth in the Capacity measure appears to capture this pattern.¹⁸

At the other end of the spectrum lie countries such as Iraq, which in 1960 possessed average scores on the Capacity measure but by 2015 had fallen far below average due to by civil and international conflict and long bouts of destructive leadership. A fourth country depicted Figure 3 – Haiti – has remained consistently low since 1990.¹⁹

Given the latent nature of the Capacity measure, we check content validity by comparing the Capacity variable with other measures that were not used in the MCMC process in order to assess whether it accurately taps the intended concept of state capacity. We choose a variety of other indicators constructed using different methodologies. If Capacity is a valid measure, we should

¹⁸The sharp decrease in Capacity in 1965-1966 is likely due to Singapore's separation from Malaysia in that period.

¹⁹The scores in Figure 3 is normalized to a 0-1 scale.

Figure 1: Posterior Distribution of Capacity in the Year 2015

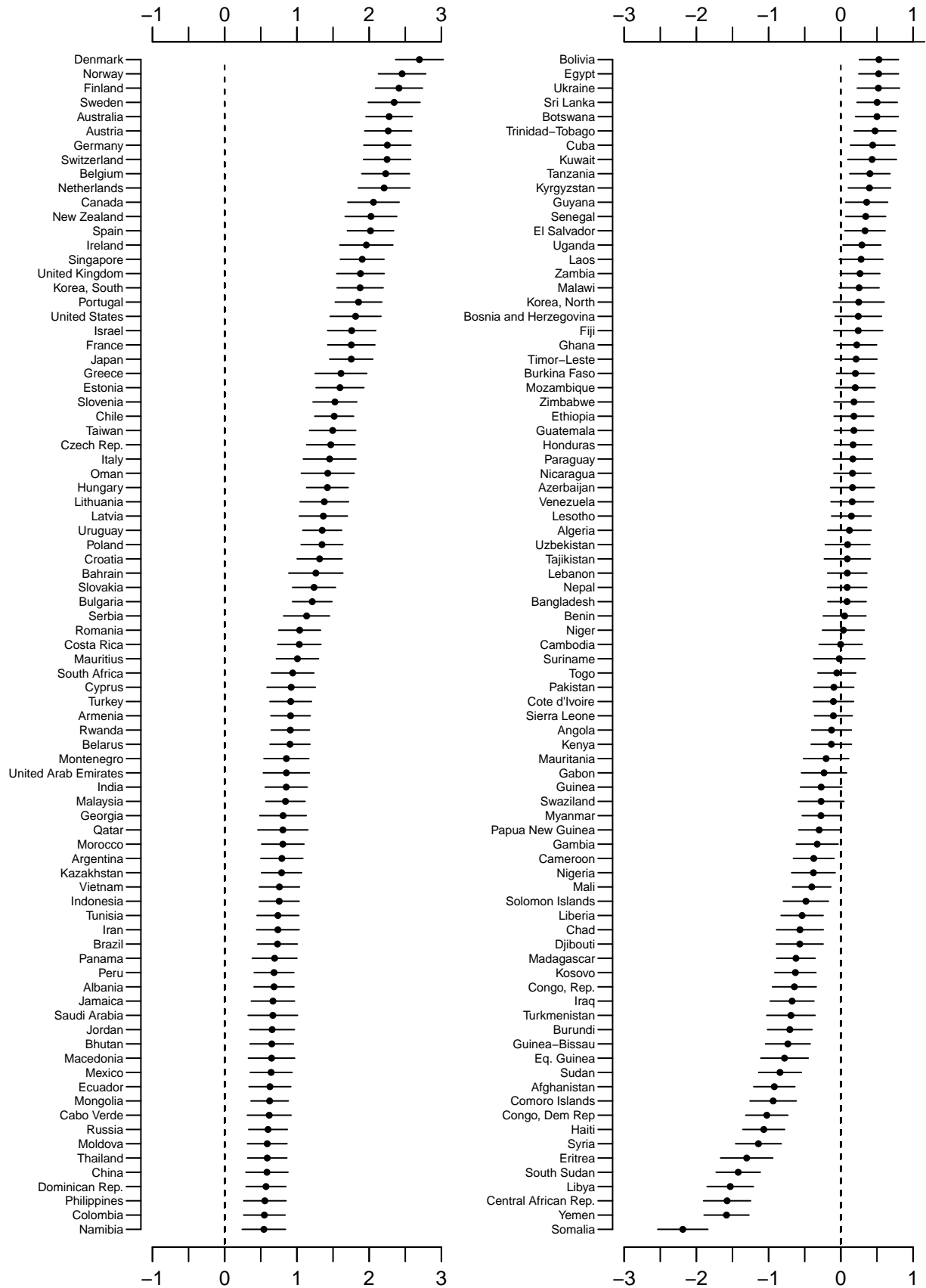


Figure 2: Scatter Plot of Capacity 1960 and 2015

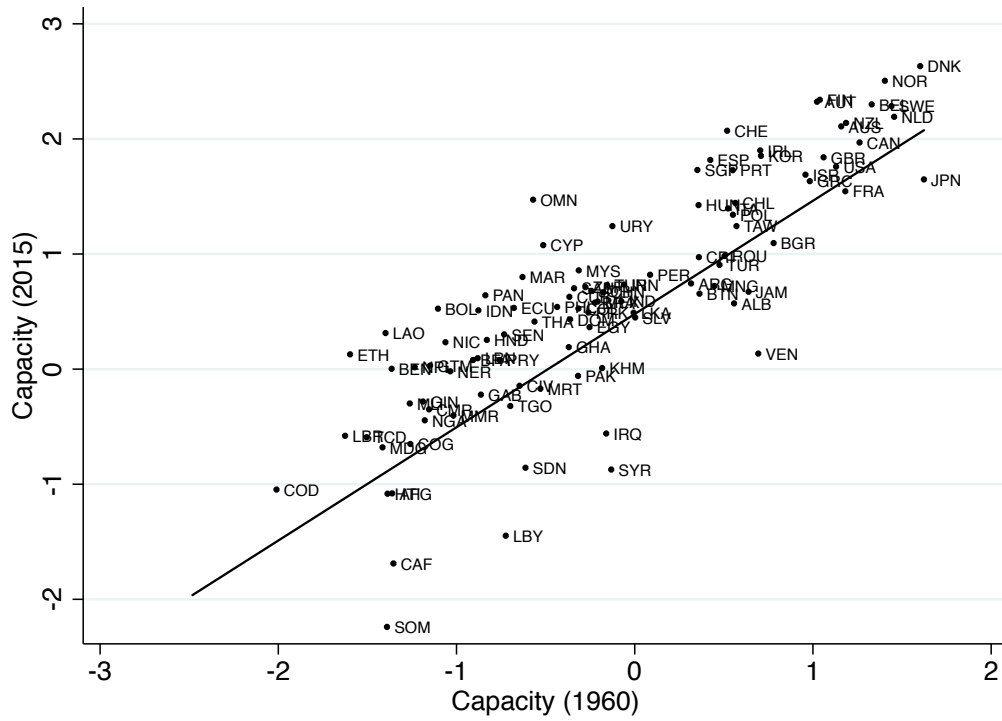
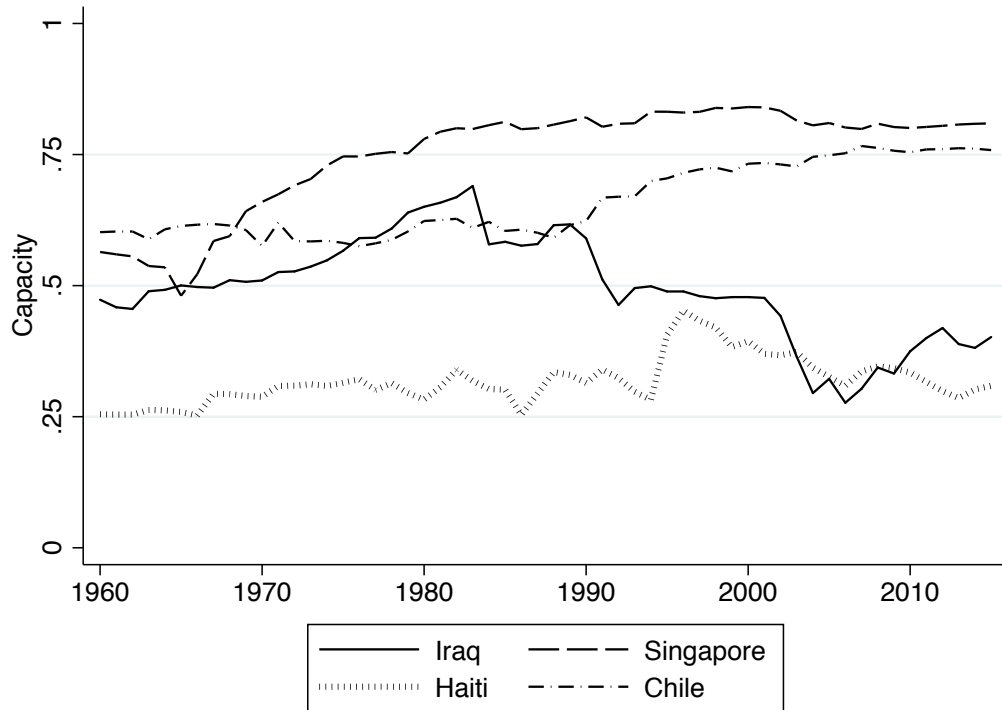


Figure 3: Capacity in Four Sample Countries (1960-2015)



observe strong correlation with other attempts to measure this concept.

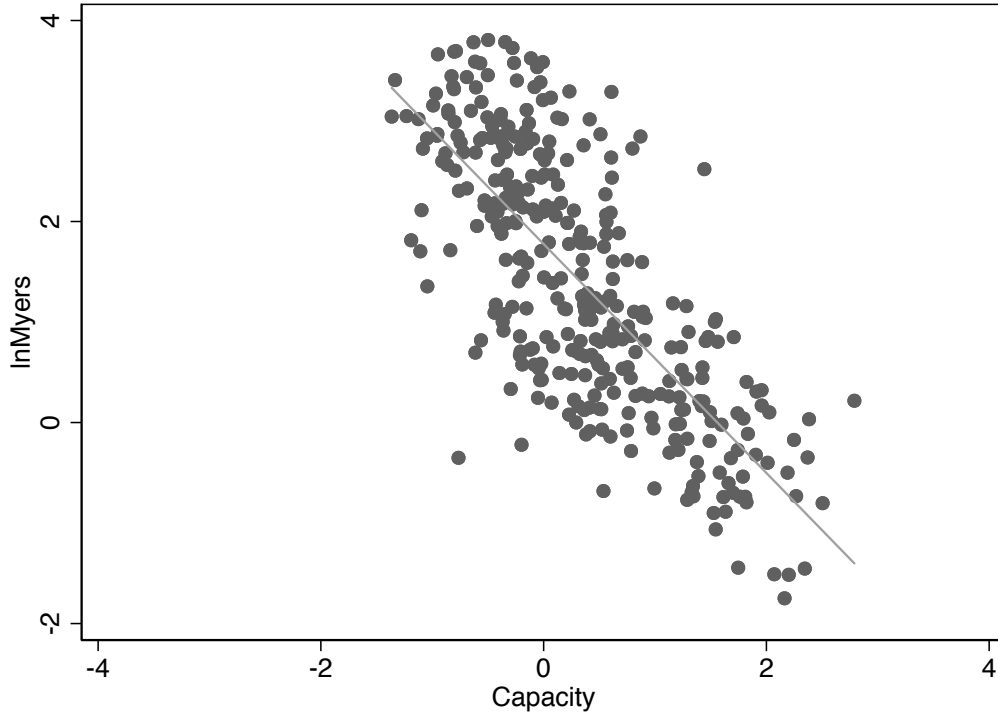
Table 3: Correlations of Estimates with Other Indicators

	Capacity	N
Government Effectiveness (WGI)	0.90	2569
Rule of Law (WGI)	0.88	2570
Basic Administration (BTI)	0.87	1247
Regulatory Quality (WGI)	0.84	2569
Rule of Law (PRS)	0.84	2569
Impartial Public Admin. (Rothstein & Teorell)	0.83	50
Public Sector Mgmt. (CPIA)	0.82	724
Rational-Legal (Hendrix)	0.82	1408
CPIA Index (World Bank Group, 2010)	0.79	724
Functioning of Government (EIU)	0.77	1296
Stateness (BTI)	0.77	1592
Functioning of Government (FH)	0.71	1642
Management Index (BTI)	0.60	1588
(log) Myers Index (Lee and Zhang)	-0.75	345
Public Services (FSI)	-0.87	1719
Fragile States Index (Rice and Patrick)	-0.88	1719

As can be seen in Table 3, the Capacity measure is quite strongly correlated in the expected direction with a broad range of other indicators in pairwise tests. Among the indicators most strongly correlated with Capacity, for example, are the WGI’s Government Effectiveness ($r = 0.90$) and Rule of Law ($r = .88$) ratings, the Fragile States Index ($r = -.88$), and BTI’s measure of Basic Administration ($r = .87$). The state-capacity indicator that is least correlated with the Capacity is BTI’s management index ($r = .60$).

For example, recent work from Lee and Zhang (2017) develops a measure of legibility – the extent of state information about citizens that is available in standardized forms – built upon the accuracy of age-reporting in national censuses. Where birth records are poor, and lack of interaction with the state creates little reason to know one’s age exactly, citizens tend report their ages to census enumerators in numbers that end with zeros or fives. The degree of “heaping” creates a way to measure legibility: the Myers index. In a dataset of 370 censuses, Lee and Zhang show a moderately strong correlation between the log Myers index and other measures of state capacity. The Capacity measure developed here is correlated more strongly with both the log Myers index ($r=-.75$) and many of those other measures than they are with each other. Figure 4 illustrates this

Figure 4: Scatter Plot of lnMyers and Capacity



relationship.

Likewise, in their book *Building State Capability*, Andrews et al. (2017) use the Public Services indicator from the Failed States Index as a measure of the state’s capability to carry out core functions. The Capacity measure is highly correlated with this measure ($r = -.87$) but has much more extensive coverage since the latter goes back only to the year 2005. Overall, given the very different measurement approaches used in this selection of other state capacity indicators, we can have more confidence that Capacity is a valid measure of the intended concept.

We further demonstrate validity in tests where we use the Capacity measure as a predictor of various outcomes widely associated with state capacity. Table 4 presents the results from six regression models that test for whether Capacity is a strong predictor of development outcomes even after controlling for log GDP per capita. In each of these tests, which involve quite different types of dependent variables, Capacity is a substantively strong and statistically significant (at the 99% level) predictor.

In Model 1, the dependent variable is a measure of the size of the shadow economy as a per-

Table 4: Construct Validity Tests for Capacity

	(1) InformalEcon	(2) lnMyers	(3) PublicServ	(4) Letters	(5) AveDays	(6) eGov
Capacity	-1.86** (0.21)	-0.54** (0.10)	-0.38** (0.09)	20.78** (4.33)	-69.80** (14.83)	0.11** (0.01)
lnGDPcap	-6.18** (0.23)	-0.57** (0.10)	-0.58** (0.10)	2.18 (2.56)	-15.87^ (8.76)	0.08** (0.01)
Constant	85.41** (1.89)	6.12** (0.77)	10.96** (0.85)	28.41 (19.62)	404.32** (67.26)	-0.22** (0.06)
N	1350	345	1719	150	150	164
R ²	0.99	0.91	0.97	0.43	0.50	0.87
Fixed Effects?	Yes	Yes	Yes	No	No	No

^ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

centage of GDP from Schneider et al. (2010).²⁰ On average across countries, they estimate the size of the shadow economy to be about 17.7%. We find that each one point increase in Capacity is associated with a reduction in the size of the shadow economy by 1.86 percentage points of GDP, controlling for the log level of GDP per capita.

Models 2 and 3 use as dependent variables the log Myers Index and the FSI Public Services indicator, discussed above. Even after controlling for log GDP per capita, Capacity is strongly associated with both of these measures. A one-unit increase in Capacity is connected with a reduction in the log Myers index to about 54% of its previous size and a decrease in the Public Services indicator by .38 points.

In Models 4 and 5, we draw upon a study conducted by Chong et al. (2014) to assess the efficiency of government in 159 countries by measuring how long it would take the country's postal service to return undeliverable mail to an international address. They sent 10 letters to each country and found that about 60% of letters were returned. The mean number of days it took to return the letter was about 228. For Model 4, the dependent variable is the percentage of letters sent to a country that were returned. Where Capacity is one point higher, the percentage of letters returned is about 20.8 percentage points higher. Similarly, in Model 5 where the dependent variable is the

²⁰The shadow economy is defined as “the shadow economy includes including “all market-based legal production of goods and services that are deliberately concealed from public authorities to avoid payment of income, value added or other taxes; to avoid payment of social security contributions; having to meet certain legal labour market standards, such as minimum wages, maximum working hours, safety standards, etc; and complying with certain administrative procedures, such as completing statistical questionnaires or administrative forms” (Schneider et al., 2010: 444).

average number of days it takes to return a letter, a one-point increase in Capacity is associated with a reduction of about 70 days in how long it takes for the letter to be returned.

Finally, Model 5 uses data from United Nations E-Government Development Database, which tracks the e-governance readiness of each UN member country's government and the extent of citizen e-participation in government. The scale runs from 0 to 1, with higher scores meaning greater preparedness. We find that each one-point increase in Capacity predicts a .11 point increase in the e-Government Development Index, which is about one-half a standard deviation in the index.

Once again, these are very different measures of state capacity and/or government performance. In each case, Capacity contributes substantial explanatory power beyond what can be explained through country wealth.

To further demonstrate the utility of the measures of state capacity created from the analysis above, we conduct two further sets of empirical tests. First, we illustrate how the broader coverage of the State Capacity Dataset expands the possibilities for research. For example, we are able to make full use of the IEG World Bank Project Performance Ratings dataset (World Bank Group, 2019), which includes ratings for over 12,000 projects completed between 1964 and 2018. These data have been used in various studies, none of which has examined the role of state capacity in the success of World Bank projects over the full span of the data. Isham et al. (1997) examine the effect of civil liberties and democracy on project success, Isham and Kaufmann (1999) study the effect of the policy environment, while Dollar and Levin (2005) assess the effect of institutional quality in the 1990s.

We define project success as a rating of either "satisfactory" or "highly satisfactory." Overall, 73.1% of projects received one of these two ratings. Figure 5 presents a Lowess curve that depicts across the values of Capacity the proportion of World Bank projects were successful. Where Capacity was higher, the rate of project success was higher as well. In Hanson and Sigman (2016), we present more sophisticated analyses which demonstrate that Capacity still has a strong effect on project success even after controlling for the level of GDP per capita, the mean years of education, and various measures of democracy and political rights.

Finally, in Table 5, we use the level of Capacity as measured in 1960, or the earliest available year

Figure 5: IEG Rating of World Bank Projects

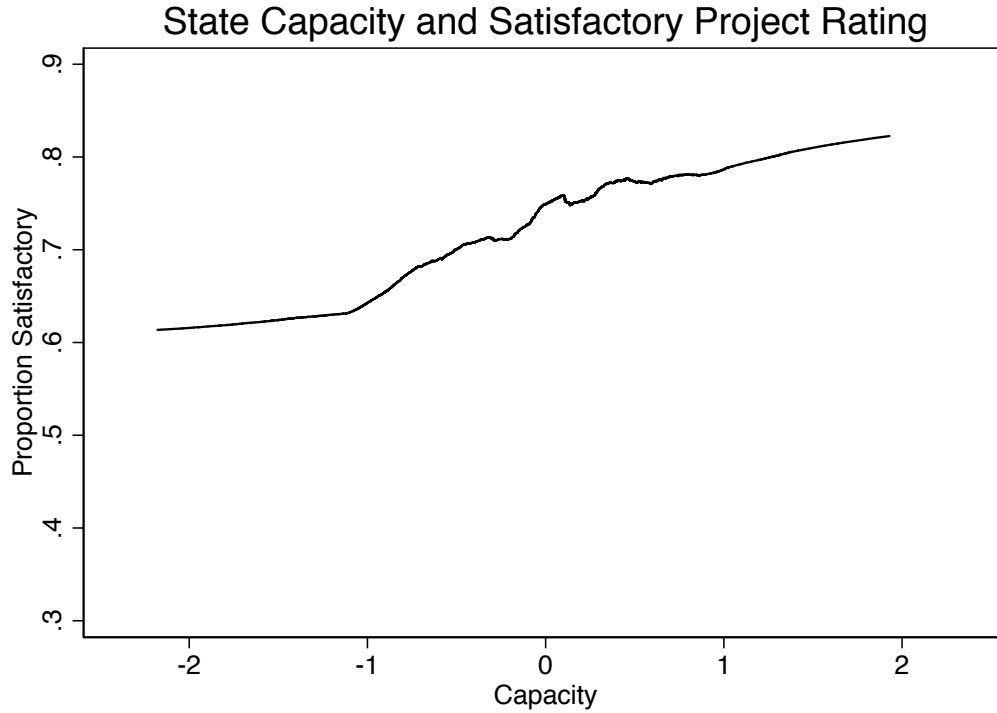


Table 5: Illustrative Tests Using Capacity

	(1) InfMort	(2) LifeExp	(3) Roads	(4) Water	(5) Hospitals	(6) lnGDPcap10
Capacity60	-15.39** (2.72)	6.43** (1.03)	43.12* (16.58)	5.06* (2.12)	1.54** (0.37)	0.58** (0.10)
lnGDPcapstart	-5.72** (1.63)	2.24** (0.62)	16.82^ (9.48)	5.49** (1.36)	0.19 (0.21)	0.61** (0.06)
Democracy	-10.34 (6.46)	4.88* (2.44)	10.78 (37.15)	13.08* (5.24)	-1.34 (0.85)	-0.03 (0.24)
TaxRev	-0.14 (0.26)	-0.11 (0.10)	-1.48 (1.65)	0.29 (0.21)	0.11** (0.04)	0.01 (0.01)
Constant	80.55** (13.65)	50.76** (5.15)	-52.81 (78.13)	28.59** (10.74)	0.71 (1.80)	3.90** (0.51)
N	148	148	106	132	112	149
R ²	0.58	0.59	0.23	0.49	0.40	0.74

^ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Cross-sectional OLS regression with standard errors in parentheses. The dependent variables, measured in the year 2010, are Roads (% of roads that are paved), Water (% of population with access to an improved water source), Hospitals (number of hospital beds per 1,000 people), InfMort (infant mortality rate), and LifeExp (level of life expectancy), and log GDP per capita. The independent variables are GDP/cap₆₀ (log level of GDP per capita in 1960), Democracy (mean level during the period 1960-2010), TaxRev (mean level of tax revenue as a percentage of GDP over the period 1960-2010), and Capacity in 1960.

for a country, as a predictor for the year-2010 levels of different health indicators, infrastructural measures, and country wealth. We consider this a very challenging test, since we control for the initial level of GDP per capita (logged), the mean level of Democracy during the period,²¹ and the mean level of tax revenues as a percentage of GDP. Each test nevertheless shows Capacity to be strongly related to these outcomes.

In Model 1, the dependent variable is a country's infant mortality rate. Where Capacity was one-unit higher in 1960, mortality in 2010 is about 15.4 deaths lower per 1,000 infants, all the other variables being held constant. Similarly, as Model 2 shows, a one unit-higher 1960 Capacity score is associated with 6.5 years longer life expectancy in 2010.

Models 3 through 5 present tests in which the dependent variables are measures of national infrastructure and health care facilities. As the estimates show, where Capacity was one unit higher in 1960, we would predict the percentage of paved roads in a country to have 43.1 points higher, the percentage of citizens with access to an improved water source to be about 5 points higher, and the number of hospital beds per 1,000 people to be 1.5 beds greater.

Finally, as Model 5 shows, Capacity in 1960 is associated with greater GDP per capita 50 years later, even after controlling for the initial level of GDP per capita. Each one-unit increase in Capacity in 1960 is associated with GDP per capita in 2010 being 58% higher.

The robustness of these results to controlling for country wealth and democracy, we argue, provides confidence that the Capacity measure is indeed capturing something that is distinct from these other concepts. Additional robustness checks are presented in the online appendix.

6 Summary

Ultimately our understanding of the causes and consequences of state capacity depends on our ability to measure it in valid, reliable, and practical ways. That state capacity is multi-dimensional, fundamentally latent, and closely related to a range of concepts presents a particularly complicated set of challenges that researchers must overcome. In focusing on the use of state capacity across

²¹Which is the polity2 index rescaled to a range of 0-1).

political science research, distilling the concept into its essential parts, identifying its underlying theoretical dimensions, and systematically analyzing the best available data for those dimensions, we hope to have advanced a nascent discussion of the conceptual and measurement issues related to state capacity, addressed recently by Centeno et al. (2017), Lindvall and Teorell (2016), Rogers (2014), Fukuyama (2013), Hendrix (2010), and Soifer (2008).

In particular, our analysis has provided new insight into the empirical manifestations of state capacity. First, we have demonstrated that it is possible to generate better geographic and temporal coverage for estimates of state capacity, thereby offering the potential to gain longer term perspectives on a range of familiar questions. Second, we have shown the difficulty of empirically disentangling the three theorized dimensions of state capacity, suggesting that researchers' intentions to isolate specific types of capacity may be difficult to achieve. Third, we have shown that using measures of state capacity, as opposed to measures that capture related institutional concepts, help to clarify long-standing debates about the relationship of institutions, human capital and economic growth.

With the rise of "big data" and sustained interest in the state as a conceptual variable in political science research, we are confident that measurement options will grow in the coming years. To make meaningful improvements on the data currently available for political science research, however, those collecting and coding data must carefully consider the issues laid out in this article, particularly as they relate to the need to focus on core functions of the state, disaggregate the state's capacities, to expand coverage of existing measures, and to eschew definitions of state capacity that relate too closely to decision-making procedures. Only then will political scientists be able to make meaningful progress assessing the effects of state institutions on a broad variety of outcomes.

References

- Adcock, Robert and David Collier. 2001. "The American Political Science Review." *Measurement validity: A shared standard for qualitative and quantitative research* 95(3): 529–546.
- Adelman, Irma and Cynthia Taft Morris. 1967. *Society, Politics, & Economic Development: A Quantitative Approach*. Baltimore: The Johns Hopkins Press.
- Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2017. *Building State Capability: Evidence, Analysis, and Action*. Oxford: Oxford University Press.
- Ang, Yuen Yuen. 2016. *How China escaped the poverty trap*. Cornell University Press.
- Arel-Bundock, Vincent and Walter Mebane. 2011. "Measurement Error, Missing Values and Latent Structure in Governance Indicators." Presented at the 2011 Annual Meeting of the American Political Science Association.
- Bäck, Hanna and Axel Hadenius. 2008. "Democracy and State Capacity: Exploring a J-Shaped Relationship." *Governance: An International Journal of Policy, Administration, and Institutions* 21(1): 1–24.
- Bates, Robert.H. 2001. *Prosperity and violence: the political economy of development*. New York: Norton.
- Berkman, Heather, Carlos G. Scartascini, Ernesto Stein, and Mariano Tommasi. 2008. "Political Institutions, State Capabilities, and Public Policy: An International Dataset." Washington, D.C., United States: IDB Research Department.
- Bersch, Katherine and Sandra Botero. 2011. "Measuring Governance: A Defense of Disaggregation." Prepared for the American Political Science Association Conference, September 1-4, 2011.
- Bertelsmann Stiftung. 2006. "Bertelsmann Transformation Index." .
- Besley, Timothy and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." *The American Economic Review* 99(4): 1218–1244.
- Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and markets: The advantage of an early start." *Journal of Economic Growth* 7(4): 347–369.
- Brambor, Thomas, Augustin Goenaga, Johannes Lindvall, and Jan Teorell. 2016. "The Lay of the Land: Information Capacity and the Modern State."
- Brautigam, Deborah. 1996. "State Capacity and Effective Governance." In Benno Ndulu and Nicolas van

- de Walle, eds., *Agenda for Africa's Economic Renewal*, chap. 2, pp. 81–108. New Brunswick: Transaction Publishers.
- Centeno, Miguel. 2002b. *Blood and Debt: War and the nation-state in Latin America*. University Park: Pennsylvania State Univ. Press.
- Centeno, Miguel A, Atul Kohli, Deborah J Yashar, and Dinsha Mistree. 2017. *States in the Developing World*. Cambridge: Cambridge University Press.
- Chaudhry, Kiren Aziz. 1997. *The Price of Wealth: economies and institutions in the Middle East*. Ithaca: Cornell University Press.
- Chong, Alberto, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2014. “Letter Grading Government Efficiency.” *Journal of the European Economic Association* 12(2): 277–299.
- Clapham, Christopher. 1996. *Africa and the international system: The politics of state survival*, vol. 50. Cambridge University Press.
- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerl, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi ting Wang, Tore Wig, , and Daniel Ziblatt. 2019. “V-Dem [Country-Year/Country-Date] Dataset v9.” Varieties of Democracy (V-Dem) Project.
- Crone, Donald K. 1988. “State, social elites, and government capacity in Southeast Asia.” *World Politics* 40(2): 252–268.
- Dahlström, Carl and Victor Lapuente. 2017. *Organizing leviathan: Politicians, bureaucrats, and the making of good government*. Cambridge University Press.
- D’Arcy, Michelle and Marina Nistotskaya. 2017. “State first, then democracy: Using cadastral records to explain governmental performance in public goods provision.” *Governance* 30(2): 193–209.
- Dincecco, Mark. 2017. *State Capacity and Economic Development: Present and Past*. Cambridge University Press.
- Dollar, David and Victoria Levin. 2005. “Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries.” World Bank Policy Research Working Paper 3524.

- Evans, Peter, Dietrich Rueschemeyer, and Theda Skocpol. 1985. *Bringing the State Back In*. Cambridge: Cambridge University Press.
- Fortin-Rittenberger, Jessica. 2014. "Exploring the Relationship Between Infrastructural and Coercive State Capacity." *Democratization* 21(7): 1244–1264.
- Fukuyama, Francis. 2013. "What is Governance?" *Governance* .
- Garretón, Manuel Antonio and Gonzalo Cáceres. 2003. "From the Disarticulation of the State to the modernization of Public Management in Chile: Administrative Reform without a State project." *Reinventing Leviathan, The Politics of Administrative Reform in Developing Countries*, Miami, University of Miami, North South Center Press. .
- Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800-2000." *American Journal of Political Science* 50(4): 893–908.
- Gerring, John. 2011. *Social science methodology: A unified framework*. Cambridge: Cambridge University Press.
- Giraudy, Agustina and Juan Pablo Luna. "Unpacking the states uneven territorial reach: evidence from Latin America."
- Golden, Miriam and Brian Min. 2013. "Distributive politics around the world." 16: 73–99.
- Hanson, Jonathan K. 2015. "Democracy and State Capacity: Complements or Substitutes?" *Studies in Comparative International Development* 50(3): 304–330.
- Hanson, Jonathan K. and Rachel Sigman. 2016. "State Capacity and World Bank Project Success." Unpublished Manuscript.
- Harbers, Imke. 2015. "Taxation and the Unequal Reach of the State: Mapping State Capacity in Ecuador." *Governance* 28(3): 373–391.
- Hendrix, Cullen. S. 2010. "Measuring State capacity: Theoretical and empirical implications for the study of civil conflict." *Journal of Peace Research* 47(3): 273–285.
- Henisz, Witold. 2000. "The Institutional Environment for Multinational Investment." *Journal of Law, Economics, and Organization* 16(2): 334–364.
- Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative lessons in authority and control*. Princeton: Princeton University Press.

- Hertog, Steffen. 2011. *Princes, brokers, and bureaucrats: Oil and the state in Saudi Arabia*. Cornell University Press.
- Howell, Llewellyn D. 2011. "International Country Risk Guide Methodology." The PRS Group, Inc.
- Huntington, Samuel. 1968. *Political Order in Changing Societies*. New Haven: Yale University Press.
- Isham, Jonathan and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *The Quarterly Journal of Economics* 114(1): 149–184.
- Isham, Jonathan, Daniel Kaufmann, and Lant H. Pritchett. 1997. "Civil Liberties, Democracy, and the Performance of Government Projects." *The World Bank Economic Review* 11(2): 219–240.
- Jackson, Robert H. and Carl G. Rosberg. 1982. "Why Africa's Weak States Persist: The Empirical and the Juridical in Statehood." *World Politics* 35(1): 1–24.
- Kaufmann, Daniel, Art Kraay, and M. Mastruzzi. 2003. "Governance matters III: Governance indicators for 1996-2002." *World Bank Policy Research Working Paper No. 3106* .
- . 2007. "The worldwide governance indicators project: answering the critics." *World Bank* <http://siteresources.worldbank.org/> .
- Kocher, Matthew. 2010. "State Capacity as a Conceptual Variable." *Yale Journal of International Affairs* .
- Kramon, Eric and Daniel N Posner. 2013. "Who benefits from distributive politics? How the outcome one studies affects the answer one gets." *Perspectives on Politics* 11(2): 461–474.
- Krasner, Stephen D. 1978. *Defending the National Interest: Raw Materials Investments and US Foreign Policy Paper*. Princeton: Princeton University Press.
- Lee, Melissa M. and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *The Journal of Politics* 79(1): 118–132.
- Lee, Sik-Yum. 2007. *Structural Equation Modeling: A Bayesian Approach*. Hoboken, NJ: Wiley.
- Levi, Margaret. 1988. *Of Rule and Revenue*. Berkeley: University of California Press.
- . 2002. "The State of the Study of the State." In Ira Katznelson and Helen V. Milner, eds., *Political science: the state of the discipline*. New York: WW Norton.
- Lieberman, Evan S. 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.

- Lindvall, Johannes and Jan Teorell. 2016. "State capacity as Power: A Conceptual Framework." *Department of Political Science, Lund University* .
- Luna, Juan Pablo and Hillel David Soifer. 2017. "Capturing Sub-National Variation in State Capacity: A Survey-Based Approach." *American Behavioral Scientist* 61(8): 887–907.
- Marshall, Monty G. and Keith Jaggers. 2009. *Polity IV Project: Dataset Users' Manual*. Center for Systemic Peace, George Mason University.
- McMann, Kelly M, Daniel Pemstein, Brigitte Seim, Jan Teorell, and Staffan I Lindberg. 2016. "Strategies of Validation: Assessing the Varieties of Democracy Corruption Data." *V-Dem Working Paper* 23.
- North, Douglass. 1981. *Structure and Change in Economic History*. New York: Norton.
- . 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.
- Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2019. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." V-Dem Working Paper No. 21. 4th edition. University of Gothenburg: Varieties of Democracy Institute.
- Plummer, Martyn. 2010. *JAGS: A Program for Analysis of Bayesian Graphical Models*. Version 2.2.0.
- . 2012. *rjags: Bayesian Graphical Models using MCMC*. Version 3.5.
- Political Risk Services. 2010. *International Country Risk Guide ICRG*. East Syracuse, New York: The PRS Group, Inc.
- Prichard, Wilson, Alex Cobham, and Andrew Goodall. 2014. "The ICTD Government Revenue Dataset." .
- Rauch, James and Peter Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75(1): 49–71.
- Rogers, Melissa Ziegler. 2014. "Income Taxation and the Validity of State Capacity Indicators." *Journal of Public Policy* 34(2): 183–206.
- Saylor, Ryan. 2013. "Concepts, Measures, and Measuring Well: An Alternative Outlook." *Sociological Methods & Research* 42(3): 354–391.
- Schneider, Friedrich, Andreas Buehn, and Claudio E. Montenegro. 2010. "Shadow Economies All over the World." The World Bank. Policy Research Working Paper No. 5356.

- Seawright, Jason and David Collier. 2014. "Rival Strategies of Validation Tools for Evaluating Measures of Democracy." *Comparative Political Studies* 47(1): 111–138.
- Sikkink, Kathryn. 1991. *Ideas and Institutions: Developmentalism in Brazil and Argentina*. Ithaca: Cornell University Press.
- Singer, J David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In *Peace, War, and Numbers*, pp. 19–48. Beverly Hills: Sage.
- Skocpol, Theda. 1985. "Bringing the State Back In: Strategies of Analysis in Current Research." In Evans et al. (1985), chap. 1, pp. 3–37.
- Sobek, David. 2010. "Masters of Their Domains: The Role of State Capacity in Civil Wars." *Journal of Peace Research* 47(3): 267–271.
- Soifer, Hillel. 2008. "State Infrastructural Power: Approaches to Conceptualization and Measurement." *Studies in Comparative International Development* 43: 231–251.
- Soifer, Hillel and Matthias vom Hau. 2008. "Unpacking the *Strength* of the State: The Utility of State Infrastructural Power." *Studies in Comparative International Development* 43(3/4): 219–230.
- Soifer, Hillel David. 2012. "Measuring state capacity in contemporary Latin America." *Revista de Ciencia Política* 32(3).
- . 2013. "State Power and the Economic Origins of Democracy." *Studies in Comparative International Development* 48(1): 1–22.
- Sudduth, Jun Koga. 2016. "Coup Risk, Coup-proofing Strategy and Leadership Survival." Working Paper. School of Government and Public Policy, University of Strathclyde.
- Taylor, Brian D. 2011. *State Building in Putin's Russia: Policing and Coercion after Communism*. Cambridge: Cambridge University Press.
- Tilly, Charles. 1990. *Coercion, Capital, and European states, AD 990-1990*. Cambridge: Blackwell.
- Transparency International. 2009. "Corruption Perceptions Index." URL <http://www.transparency.org/>.
- Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1): 201–217.
- Vreeland, James Raymond. 2008. "The Effect of Political Regime on Civil War: Unpacking Anocracy." *Journal of Conflict Resolution* 52(3): 401–425.

- Weber, Max. 1918. *Politics as a Vocation?* In *From Max Weber: Essays in Sociology*, HH Gerth and C. Wright Mills. London: Routledge.
- . 1978. *Economy and Society : An Outline of Interpretive Sociology*. Berkeley: University of California Press.
- Whitfield, Lindsay, Ole Therkildsen, Lars Buur, and Anne Mette Kjær. 2015. *The politics of African industrial policy: A comparative perspective*. Cambridge University Press.
- World Bank Group. 2010. “Country Policy and Institutional Assessment.” URL <http://data.worldbank.org/data-catalog/CPIA>.
- . 2011a. *CPIA 2011 Criteria*. The World Bank Group. URL <http://www.worldbank.org/ida/papers/CPIAcriteria2011final.pdf>.
- . 2011b. *WDI online*. Washington, DC: World Bank Group.
- . 2012. “Bulletin Board on Statistical Capacity.” URL <http://data.worldbank.org/data-catalog/bulletin-board-on-statistical-capacity>.
- . 2019. “IEG World Bank Project Performance Ratings, 1964-2018.” URL <http://data.worldbank.org/data-catalog/IEG>. Last updated on 19-Feb-2019.