

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

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

Rachel Sigman
Ph.D. Candidate in Political Science
Maxwell School of Citizenship and Public Affairs
Syracuse University

Version: September 2013

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 article 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. We use the resulting State Capacity Dataset to provide new insight into existing theories of the influence of state capacity on development and the success of World Bank projects. It is hoped that this project will provide effective guidance and tools to researchers studying the causes and consequences of state capacity.

This manuscript was formerly entitled "Measuring State Capacity: Assessing and Testing the Options." Presented at the World Bank Political Economy Brown Bag Lunch Series, March 21, 2013.

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.

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.

This article contributes to a broad range of scholarship by identifying key dimensions of state capacity and, to the extent possible, addressing these measurement challenges. We first illustrate variation in definitions of state capacity and the empirical challenges associated with different approaches. 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 24 indicators of state capacity. The resulting State Capacity Dataset provides annual estimates of the levels of state capacity along different dimensions for all countries included in the Polity dataset (Marshall and Jaggers, 2009) from 1960-2009. We then assess the new measures with empirical tests.

1 Defining State Capacity

The meaning of state capacity varies considerably across political science research. 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 at the conceptual level, data limitations frequently lead researchers to use the same empirical measures for differing concepts.

For both theoretical and practical reasons we argue that a minimalist approach to capture the essence of the concept is the most effective way to define and measure state capacity for use in a wide range of research. As a starting point, we define state capacity broadly as the ability of state institutions to effectively implement official goals (Sikkink, 1991). 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 varying relationships with social groups.

Two principles guide us in operationalizing this definition into empirical measures of state capacity fitting 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 deliberately steer away from entanglement with other concepts of interest in political science research.

Although state capacity is known to vary across policy sectors (Krasner, 1978), an approach that seeks to measure state capacity by working up from the sector or policy level faces thorny problems. First, disaggregation aimed at measuring unevenness of state capacity across different areas leads to reductionism, since policy sectors may be divided into ever smaller components. Second, unless we focus on core functions of all states, we may conflate political decisions regarding policy priorities with the capacity to implement those policies. Like Fukuyama (2004) and Soifer and vom Hau (2008), we distinguish between a government's policy choices and the ability of states to implement these policies. Third, from an empirical standpoint, measuring numerous features of the state increases data collection burden dramatically. Even if one were to collect data on a wide range of policy areas 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.

We also seek to avoid overlap with concepts such as "good governance," institutional quality, or state autonomy that are theoretically or empirically related to state capacity as defined in this

article. Following Mann (1984), we distinguish between a state’s infrastructural and despotic power: the degree to which states are capable of implementing policies is analytically separate from how policy priorities are chosen and the extent of citizen control over elites. Thus, unlike scholarship related to good governance, we avoid definitions or measures of state capacity that relate to the political organization of decision making processes.

For example, Kaufmann et al. (2003) include voice and accountability as an indicator of governance, and Grindle (1996) includes societal participation and legitimate channels of decision making (i.e. political capacity) as a dimension of state capacity. While some studies have found that higher levels of participation and accountability may produce greater state capacity (Bäck and Hadenius, 2008; Taylor, 2011), including decision making processes of this kind decreases conceptual clarity by incorporating aspects of other concepts such as democracy. As Fukuyama (2013) notes, measures that incorporate such concepts impede our ability to examine the relationship between, for example, regime type and capacity.

2 Dimensions of State Capacity

To effectively operationalize and measure state capacity requires discerning its multiple dimensions. Skocpol (1985) uses the plural term “state capacities” for this reason, noting the unevenness of the state’s ability to achieve its goals across different functions or policy areas. Likewise, as Levi writes, “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” (2002: 34). There is considerable divergence, however, on how to conceptually disaggregate state capacity.

Following the principles laid out in the previous section we seek to identify dimensions that: 1) focus to the core functions of the 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 objec-

tive 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. 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, 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, they surely are interrelated and mutually supporting. Extractive and coercive

capacities, for example, are likely prerequisites for higher levels of administrative capacity. A high level of extractive capacity requires at least some level of administrative capacity. Coercive capacity requires revenues and administrative reach into society. Moreover, all three of these dimensions require adequate information about, and control over, territories and populations. It is not surprising, then, that it can be challenging to disentangle these dimensions empirically.

3 Measurement Strategies and Challenges

Researchers use a variety of indicators to measure state capacity, but it is not always clear if indicators are relevant to the dimension of primary theoretical interest. 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.

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. These data are available for most countries from the early-1970s onwards, generally from the IMF's *Government Finance Statistics*.¹ By tapping into other sources of tax data such as the OECD's taxation statistics, coverage for additional years is possible for several countries. 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.

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) argues, the revenue sources that are most likely to capture concepts related to state capacity

¹Lieberman (2002) provides a detailed analysis of taxation data.

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).

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 any particular level of taxation, the greater the proportion of tax revenue that comes from income taxes, the higher the expected level of administrative capacity. The opposite should be true with respect to the proportion of revenue that comes from taxes on trade.

Other indicators may help measure the extent to which revenue collections accord with expectations given that countries are situated in different contexts. In the Relative Political Capacity dataset, for example, Arbetman-Rabinowitz et al. (2012) measure political extraction as the ratio of actual tax revenue relative to an expected tax yield given a country’s GDP per capita, mineral production, exports, and other factors, to provide data for at least 128 countries annually covering the years 1960-2007. Additionally, the World Bank’s (2011a) Country Policy and Institutional Assessment (CPIA) index includes a rating of the Efficiency of Revenue Mobilization for 71 countries starting in 2005, while the International Institute for Management and Development (IMD) uses survey data to rate around 50 countries per year from 1997 onward on whether tax evasion damages public finances (IMD, 2011).

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 coercive capacity.

Other types of measures of coercive capacity include the Political Terror Scale (Gibney et al., 2011) as a measure of the extent to which the state is a perpetrator of violence.² High political terror ratings are expected to indicate lower levels of state capacity (Englehart, 2009). Additionally, 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.

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.³ We also employ the measure of the artificiality of state borders (*Fractal Borders*) from Alesina et al. (2011) and the degree to which terrain is mountainous (Fearon and Laitin, 2003) to reflect the level of difficulty a state faces in reaching its population (Herbst, 2000).

²These ratings come from two sources: Amnesty International and the U.S. State Department. We use the Amnesty International ratings whenever possible and fill in missing values from the State Department ratings.

³Specifically, 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.

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. 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.⁴ 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), the Administration and Civil Service index (Global Integrity, 2012), and the Effective Implementation of Government Decisions rating (IMD, 2011), and a measure of civil service confidence from the World Values Survey (World Values Survey Association, 2009). From the World Bank's CPIA index there are two relevant ratings: Quality of Budgetary and Financial Management and

⁴There 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).

Quality of Public Administration. None of these ratings covers a broad period of time, but the combination covers significant portions of the 1960-2009 time period with at least one indicator.⁵

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.⁶ As argued in Centeno (2002b) and Soifer (2013), countries that can conduct censuses have not only the capacity to collect information exhibit higher levels of territorial reach. Finally, we include contract-intensive money, calculated from the World Development Indicators, advocated by Fortin (2010) as a measure of the state’s capacity to regulate economic exchange.

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 24 different indicators related to the three key dimensions of state capacity. The data span 50 years (1960-2009) and up to 163 countries in a given year.⁷ 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

⁵All 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.

⁶We 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.

⁷The 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.

Table 1: Indicators of State Capacity

Variable	Countries	Years	Dimension(s)
Administration and Civil Service (Global Integrity, 2012)	85	2004-2009	Administrative
Administrative Efficiency (Adelman and Morris, 1967)	69	1960-1962	Administrative
Anocracy (calculated from Polity IV)	175	1960-2009	Coercive
Bureaucratic Quality (Political Risk Services)	148	1982-2009	Administrative
Census Frequency (calculated from UN 2011)	179	1960-2009	Administrative/Extractive
Civil Service Confidence (World Values Surveys)	88	1981-2008	Administrative
Contract-Intensive Money (WDI)	172	1960-2009	Administrative
Effective Implementation of Government Decisions (IMD, 2011)	57	1998-2009	Administrative
Efficiency of Revenue Mobilization (World Bank CPIA)	74	2005-2009	Administrative/Extractive
Fractal Borders (Alesina et al., 2011)	138	1960-2009	Coercive
Military Personnel per 1,000 in population (COW)	171	1960-2009	Coercive
(Log) Military Spending Per Million in population (COW)	168	1960-2009	Coercive
Monopoly on Use of Force (Bertlesmann Transformation Index)	127	2003-2009	Coercive
(Log) Mountainous Terrain (Fearon and Laitin, 2003)	157	1960-2009	Coercive
Political Terror Scale (Gibney et al., 2011)	170	1976-2009	Coercive
Quality of Budgetary and Financial Management (World Bank CPIA)	74	2005-2009	Administrative
Quality of Public Administration (World Bank CPIA)	74	2005-2009	Administrative
Relative Political Capacity (Arbetman-Rabinowitz et al., 2011)	152	1960-2007	Extractive
Statistical Capacity (World Bank)	123	2004-2009	Administrative/Extractive
Tax Evasion not Damaging (IMD, 2011)	57	1998-2009	Extractive
Taxes on Income as % of Revenue (IMF, WDI)	152	1970-2009	Administrative/Extractive
Taxes on International Trade as % Revenue (IMF, WDI)	155	1970-2009	Administrative/Extractive
Total Tax Revenue as % GDP (IMF, WDI, OECD)	152	1960-2009	Extractive
Weberianness (Rauch and Evans, 2000)	34	1970-1990	Administrative

to identify underlying factors.⁸ 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 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 . We follow ABM's approach in modeling the means, covariances, and residual variances of the latent factors.⁹ The intercepts c_k have independent normal priors, and the disturbance terms ϵ_k have independent normal priors with mean zero and variance σ_k . Overall, then, the various observed indicators are a linear function 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.¹⁰

In traditional factor analysis, a missing indicator would cause a case to drop out of the analysis. In the Bayesian variant of factor analysis employed here, missingness is handled in two ways. First, the absence of some data may be related to state capacity itself. In such cases, we adapt the procedure employed by ABM to model the probability that an observation is missing as a function of the latent dimensions of state capacity and a country's level of GDP per capita. Since we wish to avoid use of an outcome indicator that would be so closely related to dependent variables of interest in political economy research, we use an annualized version of the State Antiquity Index

⁸We are very grateful to Arel-Bundock and Mebane for making their programming code for this method available.

⁹Specifically, $\xi_{jit} = \gamma_{jt} + \omega_{jit}$, in which the w_j 's are country-year disturbances with mean zero and a covariance matrix \mathbf{S}_t , the inverse of which has a Wishart prior. See pp. 6-7 of Arel-Bundock and Mebane (2011) for a full discussion.

¹⁰The 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.

(Bockstette et al., 2002) instead of GDP per capita. This index, referred to as StateHist, is a basic indicator of a state’s historical roots in, and sovereign control over, its present-day territory. The probability that an indicator is missing in a given country-year is modeled as:

$$\Pr(x_{kit} \text{ missing}) = \Phi \left(b_{0j} + b_{1j} \cdot \text{StateHist} + \sum_{j=1}^J b_{j+1} \xi_{jit} \right) \quad (2)$$

Missing data are thus treated as a feature of the world that can be exploited to help estimate the level of state capacity in a country. In using this technique, we maintain the procedures in ABM in which the parameters that predict the missingness of multiple indicators from the same organization are shared across the set.

Second, we do not implement the missingness procedure when cases are missing for reasons that are unrelated to state capacity. For example, the ICRG ratings of bureaucratic quality began in 1982, so we simply let the indicator drop out in prior years. The absence of these data pre-1982 does not indicate low state capacity, and including them in the missingness equations would deflate the estimates compared to years in which the indicators are present. Data on taxation, such as total tax revenue and taxes on income, required a judgment call. These data are quite sparse for years prior to 1975, when coverage expands to 96 countries.¹¹ We thus treat 1975 as the first year in which the gathering of tax revenue information by international organizations was sufficiently broad that the absence of tax revenue data is potentially informative about the dimensions of state capacity.

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 Equations 1 and 2. In addition to the distributional assumptions described above, we supply uninformative initial values of the essential parameters.¹² For purposes of 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. The effect of this identification procedure is that the

¹¹For total tax revenue, this is up from 43 countries in 1971 and just 22 countries in 1968

¹²For example, all initial values λ_{kj} and ξ_{jit} are random draws from uniform distributions that range from -10 to 10. The variances of the distributions from which the observed indicators are drawn, and thus the potential size of ϵ_{ki} , start as random draws from a uniform distribution in the range 0 to 10.

observed indicator x_k associated with this fixed parameter will set the range for the state capacity dimension ξ_j . Successive draws lead to descriptions of the posterior distributions of the remaining parameters that produce the observed indicators of state capacity.¹³

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 prominent features.

4.1 Latent Variable Analysis Results

Our analysis focuses on the posterior distributions of the one-dimensional, two-dimensional, and three-dimensional state capacity parameters (the ξ_{jit} 's) with coverage ranging from 110 to 163 countries each year from 1960-2009. To facilitate presentation, we name these estimates according to the form Capacity Jx , where J is the number of dimensions specified in the test and the lowercase letter signifies a specific dimension. Following this format, we have Capacity1, Capacity2a, Capacity2b, Capacity3a, and so forth. Since the scales of the parameter estimates are a function of the procedures used to identify the model, and thus have no substantive meaning, we normalize them to have a mean of 0 and a standard deviation of 1. Given space constraints, other estimates such as the posterior distributions of the λ_{kj} parameters are placed in the online appendix.

To understand what factors are driving the Capacity measures, 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 theorized dimensions.

Capacity1 appears to be a general-purpose measure of state capacity that draws from indicators representing all three theorized dimensions. Several of the indicators most strongly associated with

¹³As is the case in ABM, ordinal observed indicators are transformed using inverse normal transformations of their cumulative distributions.

Capacity1 are those pertaining to administrative and extractive capacity: the IMD ratings on tax evasion and implementation of government decisions; the Census Frequency indicator; ICRG's Bureaucratic Quality ratings; the Administrative Efficiency scores of Adelman and Morris (1967); the World Bank's measure of Statistical Capacity; and the CPIA indicators for Quality of Public Administration and Efficiency of Revenue Mobilization. In terms of coercive capacity, military expenditures per capita ($r = .64$) and Monopoly of Force ($r = .63$) are highly correlated as well.

The indicators with weakest correlation to Capacity1 are the geography-based measures of fractal borders and mountainous terrain, the Relative Political Capacity measure of Arbetman-Rabinowitz et al. (2011), the number of military personnel, and the measure of public confidence in the civil service from the World Values Survey. Since the geography-based measures are invariant over time, low correlation is perhaps not surprising. Low correlations with the other measures could be due to issues with construct validity or measurement error.

In the two-dimensional model, Capacity2a and Capacity2b appear to capture fairly distinct aspects of the concept: the two measures are correlated at .22 (see the supplemental material for a table of correlations between all the Capacity measures). Additionally, there are some patterns in their relationships to the various observed indicators that can help us interpret these dimensions. Indicators of administrative capacity, for example, generally correlate more strongly with Capacity2a than Capacity2b, although some of the indicators appear important in both dimensions. Additionally, we see that indicators of extractive capacity such as overall tax revenue, income tax, and levels of tax evasion correlate more strongly with Capacity2a than with Capacity2b, while measures of coercive capacity such as per capita military expenditures and per capita military personnel are more closely related to Capacity2b.

Overall, Capacity2a thus appears to represent the dimensions of extractive and administrative capacity, while Capacity2b is more aligned with coercive capacity. Despite these tendencies, the two dimensions do overlap in their relationship to some indicators. This is not surprising given the conceptual interrelationships between the three dimensions, and the fact that some indicators are designed capture multiple dimensions.

Turning to the variables resulting from the three-dimensional model, we see that Capacity3a

Table 2: Correlations of State Capacity Measures with Base Indicators

Measures	Capacity1	Capacity2a	Capacity2b	Capacity3a	Capacity3b	Capacity3c	N
Administration and Civil Service	0.47	0.44	0.25	0.35	0.32	-0.05	205
Administrative Efficiency	0.73	0.63	0.33	0.61	0.42	-0.16	199
Anocracy	0.65	0.67	0.29	0.67	0.21	0.01	7186
Bureaucratic Quality	0.77	0.74	0.51	0.72	0.35	0.10	3457
Census Frequency	0.59	0.43	0.52	0.35	0.98	-0.06	6946
Civil Service Confidence	0.01	-0.01	0.08	-0.07	0.18	-0.09	877
Contract-Intensive Money	0.63	0.55	0.47	0.53	0.38	0.11	5600
Effective Implementation of Gov. Decisions	0.81	0.75	0.49	0.71	0.05	0.11	579
Efficiency of Revenue Mobilization	0.62	0.54	0.37	0.45	0.58	-0.12	319
Fractal Borders	0.14	0.15	-0.02	0.12	-0.10	-0.02	5926
Military Expenditures	0.64	0.53	0.83	0.78	0.22	0.73	6795
Military Personnel	0.09	0.03	0.41	0.20	-0.05	0.59	7021
Monopoly on Force	0.63	0.61	0.51	0.68	0.24	0.25	487
Mountainous Terrain	-0.10	-0.16	0.00	-0.14	0.02	0.02	6827
Political Terror Scale	-0.52	-0.55	-0.21	-0.50	-0.21	0.07	4995
Quality of Budgetary and Financial Mgmt.	0.77	0.70	0.26	0.62	0.42	-0.26	319
Quality of Public Administration	0.70	0.63	0.32	0.59	0.43	-0.14	319
Relative Political Capacity	0.15	0.22	0.06	0.21	-0.02	0.06	5816
Statistical Capacity	0.71	0.65	0.46	0.57	0.49	-0.04	845
Tax Evasion not Damaging	0.83	0.76	0.61	0.77	0.00	0.21	622
Tax on Income (proportion of revenue)	0.45	0.47	0.20	0.37	0.37	-0.22	3753
Tax on Trade (proportion of revenue)	-0.66	-0.52	-0.74	-0.61	-0.25	-0.57	4005
Tax Revenue (proportion of GDP)	0.52	0.65	0.20	0.58	0.02	0.08	3940
Weberian	0.44	0.24	0.56	0.21	0.30	0.32	714

correlates quite strongly to indicators related to all three dimensions: military expenditures, control of tax evasion, Bureaucratic Quality, Implementation of Government Decisions, Monopoly on Force, Anocracy, and so forth. Compared to Capacity3a, Capacity3b has relatively high correlation with Census Frequency, Efficiency of Revenue Mobilization and Weberian bureaucracy but considerably weaker relationships with other indicators, especially those related to coercive and extractive capacities. While it is difficult to draw any strong conclusions from these findings, it appears that Capacity3b captures some basic administrative capabilities necessary for state functions, including information-gathering capacity (census) and reliable civil servants.

The third variable generated by the three-dimensional model, Capacity3c, correlates most strongly to per capita military expenditures ($r=.73$) and military personnel ($r=.59$). While these relationships are possible signs that Capacity3c represents coercive capacity, we are conscious of the caveat mentioned earlier that states with high levels of military spending and personnel may be those engaged in conflict.¹⁴ The fact that Capacity3c is negatively correlated with indicators such as Quality of Budgetary and Financial Management, the proportion of tax revenues coming from income taxes, Administrative Efficiency, Quality of Public Administration, and Efficiency of Revenue Mobilization raises questions about whether or not the type of coercive capacity reflected in Capacity3c actually weakens administrative and extractive capacity.

Looking across the full set of indicators, the overall degree of correlation between the indicators and the estimated latent parameters declines moving from Capacity3a to Capacity 3b to Capacity3c. For example, the mean of the absolute values of the 24 correlation coefficients drops from .47 to .27 to .18 across the three measures. As with traditional factor analysis, each additional factor captures a smaller portion of remaining variation of the indicators.¹⁵ Since the most dominant factor (Capacity3a) reflects a mixture of all three dimensions, moreover, it appears difficult to empirically disentangling the three theorized dimensions.

¹⁴Using data on war incidence from the Correlates of War dataset (Singer et al., 1972) and Fearon and Laitin (2003), we find mixed evidence on the question of whether countries with higher levels of Capacity3c are more likely to be engaged in interstate and/or civil wars. Results not reported here.

¹⁵There is not a simple counterpart to the Kaiser-Guttman criterion which states that factors with eigenvalues smaller than one can be excluded.

5 Exploring the New Measures

With 7,186 observations in total, the State Capacity Dataset is much broader in scope than the most commonly used 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 broader coverage of countries and years is welcome, provided that the measures perform well. The main goal of this section is to determine whether the measures behave in the expected manner, and whether they will be useful for investigating theoretical questions regarding state capacity. We begin by examining the relationship between the Capacity estimates and indicators related to state capacity that were not used in the estimation routines. If the measures are valid, they should relate to other measures of state capacity.

Table 3: Correlations of Estimates with Other Indicators

	Capacity1	Capacity2a	Capacity2b	N
Basic Administration (BTI)	0.76	0.72	0.58	487
Control of Corruption (WGI)	0.82	0.78	0.61	1764
Corruption Perceptions Index (TI)	0.83	0.79	0.62	1639
Corruption (PRS)	0.65	0.64	0.35	3457
CPIA Index (World Bank Group, 2010)	0.82	0.75	0.27	319
Failed States Index (Rice and Patrick)	-0.81	-0.76	-0.64	866
Functioning of Government (FH)	0.72	0.71	0.36	647
Functioning of Government (EIU)	0.76	0.71	0.49	320
Good Government (Knack and Keefer)	0.69	0.64	0.63	155
Government Effectiveness (WGI)	0.88	0.85	0.64	1764
Impartial Public Admin. (Rothstein & Teorell)	0.82	0.76	0.61	50
Management Index (BTI)	0.65	0.64	0.26	357
Public Sector Mgmt. (CPIA)	0.81	0.74	0.30	319
Rational-Legal (Hendrix)	0.82	0.79	0.58	1408
Regulatory Quality (WGI)	0.85	0.82	0.57	1764
Road Density	0.37	0.36	0.23	3508
Rule of Law (PRS)	0.63	0.60	0.48	3457
Rule of Law (WGI)	0.85	0.81	0.62	1765
Stateness (BTI)	0.63	0.62	0.45	828
Voice and Accountability (WGI)	0.62	0.62	0.50	1765

As can be seen in Table 3, the Capacity1 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 Capacity1, for example, are the WGI's Government Effectiveness rating, Transparency International's Corruption Perceptions Index, the Impartial Public Administration rating from Rothstein and Teorell (2008), the World Bank's CPIA Public Sector Management score, and the Rational-Legal dimension of state capacity from Hendrix (2010). The indicator that is least correlated with Capacity1 is the Road Density measure used by Herbst (2000).¹⁶

In the two-dimensional model, the correlations exhibit a very similar pattern: indicators that correlate the most strongly with Capacity1 also correlate the most strongly with the Capacity2 measures. Consistently for each indicator, Capacity2b shows weaker correlation than Capacity2a. This result, which is also present in the three-dimensional model (see supplementary materials), accords with the analysis above that the additional dimensions are secondary in importance but help capture some common variation in the indicators that the first dimension does not capture.

5.1 Variation Across Space and Over Time

As a second check on the face validity of the new measures, we examine the relative scores for all countries in a single year and over time. The Figure 1 indicates the mean and a two-standard deviation range for each country's Capacity1 distribution, ranked from the highest to the lowest on the measure. In general, 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 has one of the highest scores helps us to know that these measures do not capture concepts more closely related to democratic governance than to capacity itself. Plots similar to Figure 1 for all the Capacity measures and a broader range of years are included in the supplemental materials.

One of the major challenges of measuring state capacity is the ability to assess changes over time. Though state institutions are generally considered stable, it is important for researchers in some fields to assess changes in state capacity. That our data provide continuous coverage from 1960 through 2009 is a big improvement over most existing measures, but it is nonetheless necessary

¹⁶The same is true for the relationship of Relative Political Reach to all the Capacity measures.

Figure 1: Posterior Distribution of Capacity1 in the Year 2000

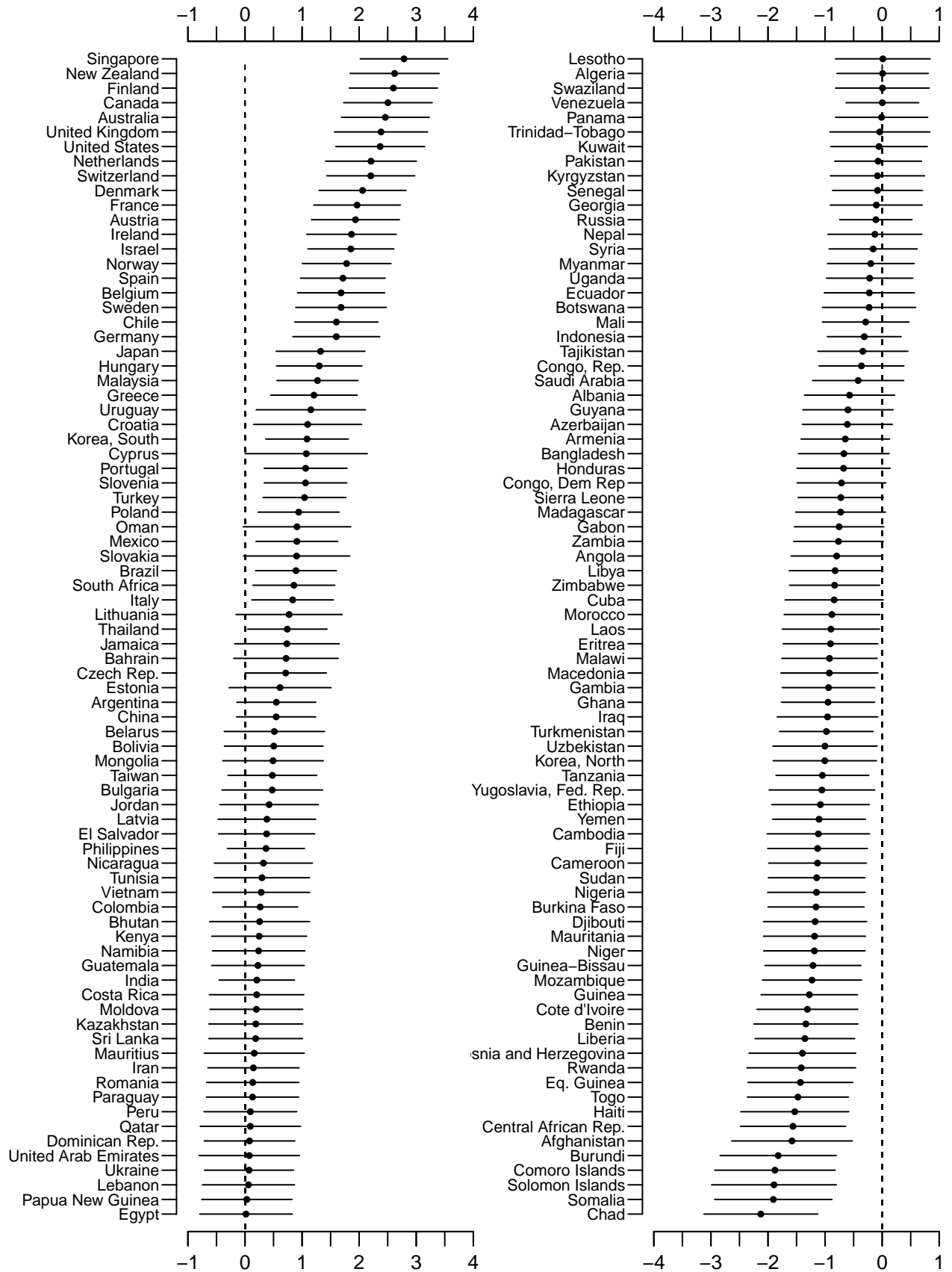
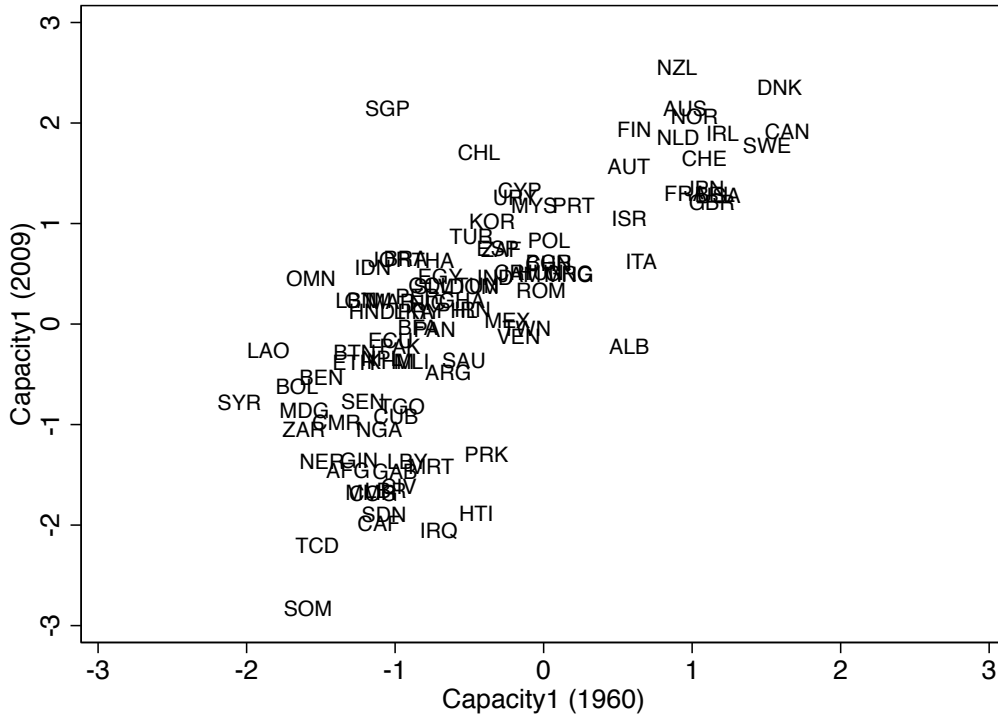


Figure 2: Scatter Plot of Capacity1 1960 and 2009

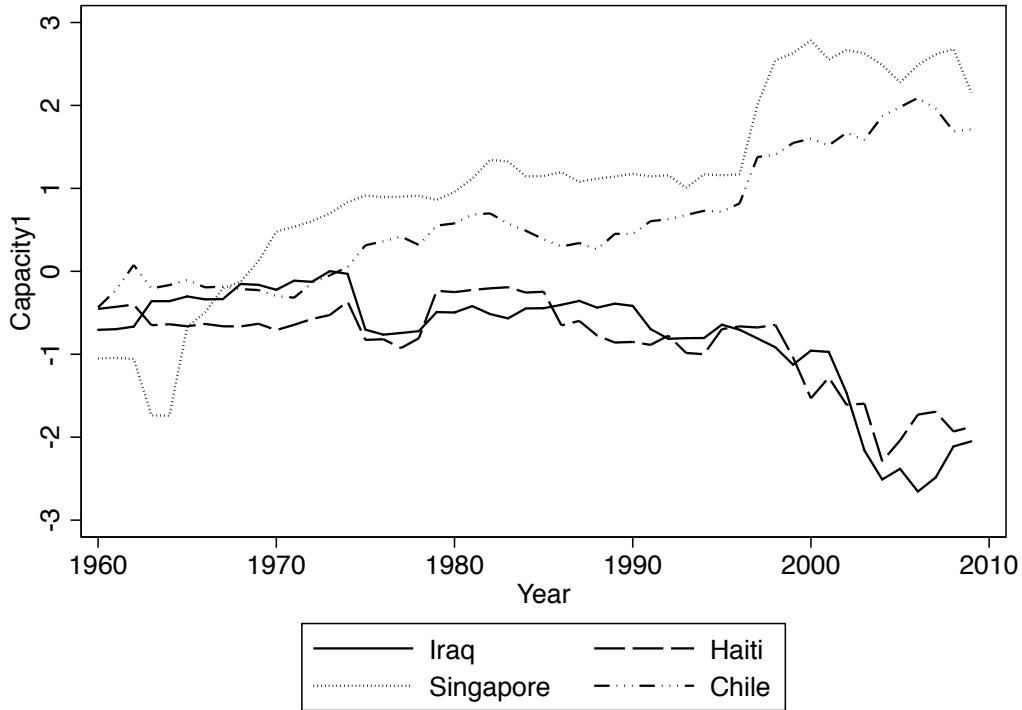


to assess the validity of the time-series aspect of our data. We seek to provide some insight into the temporal variation by first plotting scores for our measures over different years, and second with brief inquiries into several specific cases.

Figure 2 plots Capacity1 scores for all countries in the dataset, with 1960 scores on the x-axis and 2009 scores on the y-axis. As theory would predict, the relationship between the Capacity1 variables in different years is strongly positive. Most countries starting with high scores in 1960 also have high scores in 2009. 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 Capacity1 rose the most were Singapore, Chile, Oman, Indonesia, Jordan, New Zealand, and Laos. Countries where Capacity1 decreased the most include Haiti, Iraq, Somalia, North Korea, the Central African Republic, Sudan, Albania, and Chad.

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 struc-

Figure 3: Capacity1 in Four Sample Countries (1960-2009)



tures. 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 both Capacity1 appears to capture this pattern.¹⁷

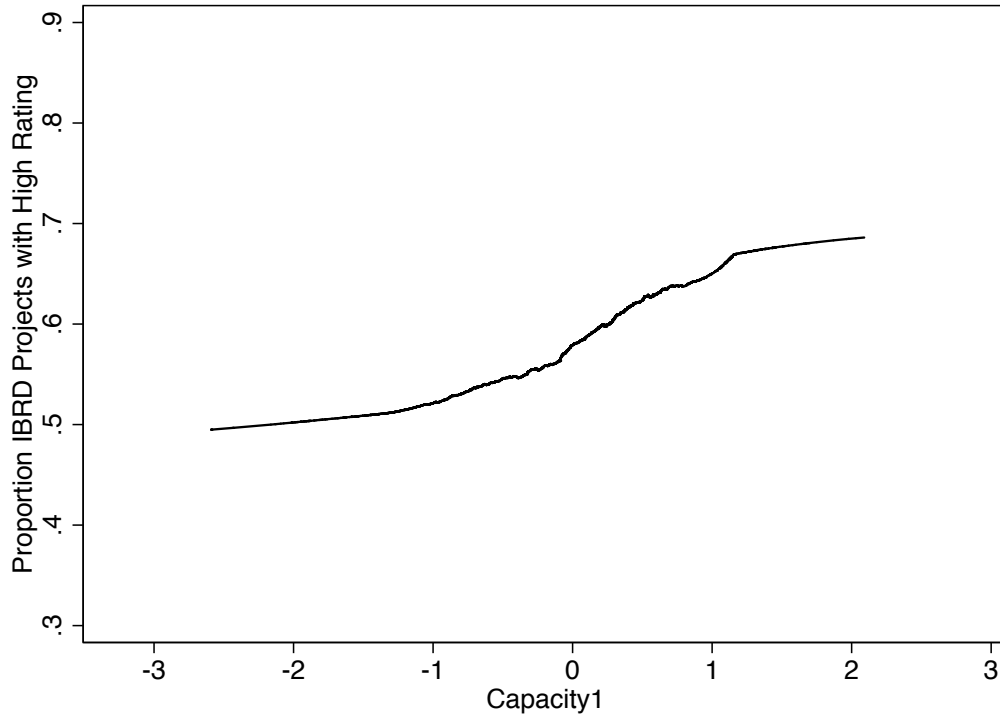
At the other end of the spectrum lie countries such as Haiti and Iraq, which in 1960 possessed average scores on the Capacity1 measure but by 2009 had fallen far below average. Torn apart by civil or international conflict and long bouts of destructive leadership, these countries have, as Figure 3 shows, seen state capacity erode since 1990.

5.2 Empirical Tests using the New Measures

To demonstrate the utility of the measures of state capacity created from the analysis above, we conduct two 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

¹⁷The sharp decrease in Capacity1 in 1965-1966 is likely due to Singapore’s separation from Malaysia in that period.

Figure 4: IEG Project Success Rating Across Values of Capacity1



of the IEG World Bank Project Performance Ratings dataset (World Bank Group, 2013), which includes ratings for 9,000 projects completed between 1964 and 2011. 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, 58.3% of projects received one of these two ratings. Figure 4 presents a Lowess curve that depicts across the values of Capacity1 the proportion of World Bank projects were successful. Where Capacity1 was higher, the rate of project success was higher as well. In the supplemental materials, we present more sophisticated analyses which demonstrate that Capacity1 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.

In our second empirical test, we re-examine the finding in Glaeser et al. (2004) that high levels

of human capital lead to improvement in institutions but that institutions have no effect on the improvement of human capital.¹⁸ The original test contains two parts. First, one set of tests shows that measures of institutional quality – specifically constraints on rulers and measures of democracy or autocracy – have no effect on improvement in the level of human capital over the subsequent five years. To measure improvement in human capital, they use the five-year change in the average number of years of schooling (*YearsEduc*) in the population. A second set of tests shows that higher levels of human capital are linked with improvement in institutions over the next five years.

We replicate these tests using our measures of Capacity1, Capacity2a, and Capacity2b in Table 4.¹⁹ In Panel A, in contrast to Glaeser et al., the state capacity measures are associated with improvement in human capital in three of the four tests. For example, an increase in Capacity1 by one unit is associated with an increase in the mean years of schooling by about 0.12 over the next five years (Model 1). Models 3 and 4 show similar effects for the Capacity2b measure and the combined Capacity2 measures. Only when Capacity2a is entered by itself (Model 2) do we observe no discernible effect of state capacity on growth of human capital.

Unlike Glaeser et al. (2004), furthermore, we find no systematic relationship between the level of human capital and improvement in our respective measures of institutional quality over the next five years. As can be seen in Panel B, the level of human capital is found to have no effect on the level of Capacity1, a negative effect on the level of Capacity2a, and a positive effect on Capacity2b. Overall, then, state capacity is much more consistently linked to subsequent growth in human capital than the reverse.

We thus dispute the conclusion in Glaeser et al. that countries first accumulate human and physical capital and then “improve their institutions” (2004: 298). This conception of institutions is too strongly linked to regime type rather than the capability of the state, and it leads to an incorrect inference about institutions writ large. Once we employ our measures of state capacity, we restore the link between institutions and variation in the growth of human capital across countries and thus reassert the value of looking for the “deep” factors that explain economic growth.

¹⁸Specifically, this finding is presented in Table 12 of Glaeser et al. (2004).

¹⁹Both our replication with the original variables and with the Capacity3 measures are included in the online appendix.

Table 4: Replication of Glaeser et al. (2004) Table 12 with New Measures

	(1) Capacity1	(2) Capacity2a	(3) Capacity2b	(4) Capacity2
<i>Panel A: Dependent variable is the 5-year change in years of schooling ($t+5, t$)</i>				
Years of Schooling _{<i>t</i>}	-0.09** (0.02)	-0.06** (0.02)	-0.09** (0.02)	-0.09** (0.02)
Log GDP per capita _{<i>t</i>}	0.13* (0.06)	0.18** (0.06)	0.15* (0.06)	0.11 (0.07)
Capacity1 _{<i>t</i>}	0.12** (0.04)			
Capacity2a _{<i>t</i>}		0.00 (0.03)		0.05 [^] (0.03)
Capacity2b _{<i>t</i>}			0.09** (0.03)	0.12** (0.03)
Constant	-0.26 (0.43)	-0.80 [^] (0.46)	-0.38 (0.46)	0.00 (0.52)
N	634	634	634	634
R ²	0.06	0.04	0.06	0.06
<i>Panel B: Dependent variables are the 5-year changes in state capacity ($t+5, t$)</i>				
Years of Schooling _{<i>t</i>}	-0.00 (0.02)	-0.08** (0.03)	0.18** (0.02)	
Log GDP per capita _{<i>t</i>}	0.34** (0.09)	0.41** (0.10)	0.13 [^] (0.07)	
Capacity1 _{<i>t</i>}	-0.59** (0.06)			
Capacity2a _{<i>t</i>}		-0.64** (0.06)		
Capacity2b _{<i>t</i>}			-0.65** (0.05)	
Constant	-2.51** (0.64)	-2.75** (0.71)	-1.98** (0.60)	
N	716	716	716	
R ²	0.32	0.33	0.34	

Notes: This table replicates Table 12 in Glaeser et al. (2004), replacing the original measures of institutions with the state capacity measures. The estimation method is OLS regressions with country fixed effects for the cross-section of countries. Errors are clustered at the country level and reported in parentheses.

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

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 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 Soifer (2008), Hendrix (2010) and Fukuyama (2013).

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

- Adelman, Irma and Cynthia Taft Morris. 1967. *Society, Politics, & Economic Development: A Quantitative Approach*. Baltimore: The Johns Hopkins Press.
- Alesina, Alberto, William Easterly, and Janina Matuszeski. 2011. "Artificial States." *Journal of the European Economic Association* 9(2): 246–277.
- Arbetman-Rabinowitz, Marina, Jacek Kugler, Mark Abdollahian, Kristin Johnson, and Kyungkook Kang. 2011. "Replication data for: Relative Political Capacity Dataset." URL <http://hdl.handle.net/1902.1/16845>. Transresearch Consortium [Distributor] V3 [Version].
- . 2012. *Performance of Nations*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- 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." .
- 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.
- Centeno, Miguel. 2002b. *Blood and Debt: War and the nation-state in Latin America*. University Park: Pennsylvania State Univ. Press.
- Crone, Donald K. 1988. "State, social elites, and government capacity in Southeast Asia." *World Politics* 40(2): 252–268.

- Dollar, David and Victoria Levin. 2005. "Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries." World Bank Policy Research Working Paper 3524.
- Englehart, N. A. 2009. "State Capacity, State Failure, and Human Rights." *Journal of Peace Research* 46(2): 163–180.
- Evans, Peter, Dietrich Rueschemeyer, and Theda Skocpol. 1985. *Bringing the State Back In*. Cambridge: Cambridge University Press.
- Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(01): 75–90.
- Fortin, Jessica. 2010. "A Tool to Evaluate State Capacity in Post-Communist countries, 1989–2006." *European Journal of Political Research* 49: 654–686.
- Fukuyama, Francis. 2004. "The Imperative of State-Building." *Journal of Democracy* 15(2): 17–31.
- . 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.
- Gibney, Mark, Linda Cornett, and Reed Wood. 2011. "Political Terror Scale, 1976-2008." URL <http://www.politicalterror scale.org/>. Retrieved on April 23, 2013 from the Political Terror Scale Web site.
- Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3): 271–303.
- Global Integrity. 2012. *Global Integrity Report*.
- Grindle, Merilee S. 1996. *Challenging the State: Crisis and Innovation in Latin America and Africa*. Cambridge: Cambridge University Press.
- 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.
- IMD. 2011. *World Competitiveness Online*.
- 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.
- 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* .
- Krasner, Stephen D. 1978. *Defending the National Interest: Raw Materials Investments and US Foreign Policy Paper*. Princeton: Princeton University Press.
- 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.
- Mann, Michael. 1984. "The autonomous power of the state: its origins, mechanisms and results." *European Journal of Sociology* 25(02): 185–213.
- Marshall, Monty G. and Keith Jagers. 2009. *Polity IV Project: Dataset Users' Manual*. Center for Systemic Peace, George Mason University.
- North, Douglass. 1981. *Structure and Change in Economic History*. New York: Norton.
- 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.
- Rauch, James and Peter Evans. 2000. “Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries.” *Journal of Public Economics* 75(1): 49–71.
- Rothstein, Bo and Jan Teorell. 2008. “What is Quality of Government? A Theory of Impartial Government Institutions.” *Governance* 21(2): 165–190.
- 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.
- 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. 2013. “State Power and the Economic Origins of Democracy.” *Studies in Comparative International Development* 48(1): 1–22.
- 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.
- 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.
- . 2013. “IEG World Bank Project Performance Ratings, 1964-2013.” URL <http://data.worldbank.org/data-catalog/IEG>. Last updated on 01-Feb-2013.
- World Values Survey Association. 2009. “World Values Survey 1981-2008 Official Aggregate.” V.20090901. Aggregate File Producer: ASEP/JDS, Madrid. (www.worldvaluessurvey.org).