Laying the Foundations of the State: Identifying the Constituent Factors of State Building

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

Version: April 2019

Abstract

This paper seeks to synthesize theories from the comparative politics and international relations literature about what leads to state building and the growth of state capacity. It draws upon these theories to develop a range of empirical propositions and test them using the State Capacity Dataset (Hanson and Sigman 2013). The objective is for these findings to help determine the relative importance of different explanatory factors – such as geography/terrain, social diversity, state rivalries, colonial history, politcal regime type, and international assistance – and facilitate the process of bringing them together into a holistic framework.

Prepared for delivery at "State Capacity. Concepts – Causes – Consequences." Duke Workshops in the Frontiers of Political Science. Duke University. April 12-13, 2019.

Across the comparative politics and international relations literature, a range of theories seeks to explain why states form and become strong in some cases, while remaining weak and fragile in other cases. Rather than add to this collection of theories, this paper draws upon them and develops a set of empirical propositions with the goal of determining just how these various ideas best fit together. Leveraging a new data set that measures the level of state capacity annually from 1960 to 2015 for most countries in the world, this paper is able to bring new evidence to bear that can help determine the relative importance of different explanatory factors.

1 Theories of State Formation

Typically, scholarship on state formation is historical in nature and focuses for the most part on the development of states in the European context. More recent work, such as Herbst (2014), broadens the scope of these theories by adapting them to other parts of the world, changing critical assumptions to reflect historical conditions in these contexts more accurately. Due to the fertility of this scholarship, we have no shortage of theories. More likely, we can make greater progress in understanding what factors lead to the formation of strong states by putting these theories to the test empirically.

In this section, I assess major theories with an eye toward explaining changes in the levels of state capacity across countries in the 1960-2015 time period. By empirically testing predictions derived from these predications using an updated version of the *State Capacity Dataset* from Hanson and Sigman (2013), I hope to produce findings that apply not only to this comparatively data-rich period of time but also to earlier historical periods as well.

The early literature on state formation describes a process that transpires over centuries. Human communities develop impersonal, permanent political institutions that acquire governing authority. Over time, these institutions can also acquire moral authority and gain the allegiance of individuals over localities, kinship-based units, and religious organizations (Strayer 1970). State formation involves the development of

systems for the administration of justice, the collection of revenues, and the provision of security. Successful states not only become the unquestioned authority in a particular territory but also acquire legitimacy in the eyes of those who dwell within those boundaries. Furthermore, states must secure a steady stream of revenues extracted from the territory under their control.

This process can be path dependent. The legacy of a state's development serves to set the context for subsequent possibilities. For example, the success of previous efforts to build state capacity and enhance the legitimacy of the state may create conditions favorable for additional expansions in response to emerging needs, while the failures of the past may leave in place a stultifying political-institutional constellation in which states do not adapt to changing circumstances. In this scenario, states remain weak, fail, or are defeated by stronger neighbors.

In broad terms, there are three types of factors that affect the process of state formation and growth of state capacity. First, there can be incentives for capacity-building that arise from internal factors related to governance of the territory. Rulers may build greater state capacities in order to solidify their rule, increase their revenues, facilitate internal commerce, defend the interests of dominant classes, and settle disputes. Second, external factors affect the extent to which rulers must concern themselves with threats from outside the territory controlled by the state, or with the opportunity for gains by expanding this territory. Third, there are a range of geographical and demographic conditions that affect the ease with which the state can establish institutional roots in an area, project power throughout the territory, and regulate society.

Scholars in comparative politics tend to focus on the first of these three categories, with some attention to the third. States emerge as the legal-institutional expression of dominant actors or social forces, and they evolve in their capabilities over time as a function of the policies they are tasked to perform by state actors. Yet, as these institutions become entrenched, and perhaps gain legitimacy within the populace, they also begin to take on a life of their own and serve to help define the incentives facing these

actors and the options available to them. The endogenous nature of the relationship between institutions and political actors is described well in North (1990).

International relations scholars, naturally, are more focused on the state as an actor in the international system, particularly with respect to the role of conflict. State formation and the construction of state capacities are responses to external threats and changes in military technology (Spruyt 2011). The bellicist theory of state formation sees state-building as an effort to build the coercive and extractive capabilities of the state to to defend against external threats (Tilly 1990). War serves as a source of selective pressure to weed out weaker states.

Most likely, the reality is that all of these factors matter, and we should seek to integrate them in a coherent fashion. State capacities evolve as state actors respond to incentives to implement particular policies or defend against threats, and citizens accord legitimacy to the state to the degree that they accept this authority. Whether the motivation for these actions arises from internal or external conditions is not the fundamental issue. They are neither mutually exclusive nor competing explanations. Instead, we can treat all of these factors as creating a range of possible stimuli for state formation and capacity-building that exist to varying degrees in different times and places. Empirically, the question becomes one of whether particular factors have the expected effects when they are present.

1.1 The Provision of Security

With this argument in mind, I now examine several prominent explanations to identify the mechanisms through which they are expected to affect state formation and the growth of state capacity. The first among these is the bellicist perspective. As Hintze (1975b) argues, in the beginning all state organization was military organization. Citizens banded together either in defense of territory or for purposes of seizing territory held by others. As the nature of threats grew and evolved due to changes in military technology and organization, the capacity of the state to wage war had to expand

in response, leading to the emergence of territorial nation-states with large standing armies.

Wars can affect state development in several ways. They create pressure to develop the infrastructure to defend territory. When borders are contested, states must do more to extend their coercive reach out to these frontier areas, creating stronger linkages between the centers of power and outlying areas. Wars also force states to develop extractive capacities to mobilize manpower and money, requiring the development of bureaucratic organization (Tilly 1990). Additionally, wars serve to create nationalist sentiments that can build a sense of citizenship and cohesion (Hintze 1975a; Dannreuther 2007). One theorized mechanism for this effect is the large-scale mobilization of troops to fight in international conflict, superseding local or regional identities with national ones.

Tilly (1990) argues that state development in response to war has coercion-intensive and capital-intensive dimensions. The former requires the development of an extractive structure to squeeze revenues and manpower out of the population; the latter involves access to revenues through bargains with capitalists that can provide the resources to fund military operations. The strongest states have capacity in both dimensions. An important component of his theory is that competition between states creates the need to develop these capacities or face defeat. Rivalries thus foster stronger states.

The dominant perspective from international relations literature, however, is that state development in the modern era is fundamentally different from earlier times due to changes in the nature of the international system. First, territorially-based sovereign states superseded other forms of organizing the international system and, in particular, the organization of extraterritorial violence (Spruyt 1994; Thomson 1994). Second, borders between states are for the most part established, recognized internationally, and often defended by dominant powers in the international system when disputes arise. In earlier times, frontier areas stood between different centers of power, leaving the precise boundaries subject to contestation. States in the post-WW II era less

frequently face the same security imperatives that led to the extension of state infrastructure out to the territorial boundaries.

1.2 The Creation of a Political Community

Other theories focus more on internal factors that relate to the construction of a political community that confers legitimacy to the state. Holsti, for example, emphasizes the importance of the "ideas and myths that sustain the legitimacy of political orders and the communities on which those orders are based" (1996, 45). In Holsti's framework, legitimacy has both horizontal and vertical components. The former refers to agreement over the definition of the political community; the latter pertains to the principles on which the right-to-rule is based. This framework serves as a useful tool.

Baseline conditions for the construction of horizontal legitimacy differ. The greater the heterogeneity of the population, the less likely that there exists a shared sense of community. Loyalties to locality, kinship-based unit, or religion may dominate. Thus, where ethnic, linguistic, or religious diversity is greater, state builders face greater challenges in creating the idea that all people in the territory should be ruled by a common political order. As Weber's (1976) study of France makes clear, these challenges are not insurmountable, but they do require significant effort and time. Not until the 20th century did the idea of France as a nation take full hold. Prior to that time, inhabitants "knew themselves to be French subjects, but to many this status was no more than an abstraction" (p. 486). What changed was the growth of transportation linkages, broader markets, greater uniformity of schooling, and military service that created a shared understanding of what it meant to be French.

Challenges of horizontal legitimacy may be particularly acute when groups with a shared identity are split by territorial borders. Colonialism left in place a system of states that were created through processes that reflected the interests of the colonizing powers, producing borders that often bore little correspondence to population patterns. Generally, these borders were accepted by rulers in the post-colonial era (Herbst

2014), but this fact did not make the task of creating political community in the newly-independent states easier. As Holsti argues, a political community requires not just a notion of citizenship but a sentiment that binds people together. Accordingly, one effect of a colonial legacy is to temper the strength of political community.

Vertical legitimacy requires a widespread acceptance of the regime's right to rule. In earlier times, this authority may be based upon the divine rights of kings or the mandate of heaven. The Enlightenment era brought a fundamental shift in values that made the "consent of the governed" the primary source of legitimacy, led to the creation of liberal democracies, and influenced absolutist rulers to reframe themselves as servants of the people. Although history provides examples of autocrats that brought considerable improvements to the lives of common citizens, thus sustaining the legitimacy of their regimes, democratic regimes are logically much more conducive to the creation of vertical legitimacy.

Ertman (1997) argues for the importance of strong, representative assemblies in affecting the character of state development. These assemblies produced constitutional, rather than absolutist, states with greater internal cooperation and stronger connections between central and local governments. When participatory politics at the local level are combined with a strong political center, conditions are optimal for balanced political, social and economic development. This combination prevents the rise of patrimonialism in politics and fosters the rise of the modern bureaucratic state.

1.3 The Revenue Imperative

Since revenue collection is so fundamental to the operation of the state, the relationship between the nature of revenue collection and the state's development is naturally very strong. As Levi succinctly puts it, "[t]he history of state revenue production is the history of the evolution of the state" (1988, 1). The form and extent of taxation in Levi's formulation is a function of the bargaining power of the ruler relative to subjects and the costs of revenue collection. Actions of rulers to maximize revenues subject to

these constraints leads to the evolution of state structures. Capable, effective states have reliable sources of revenue.

Revenue collection varies in difficulty across contexts according to many factors: the nature of production, the degree of commercial activity, the extent of compliance on the part of constituents, the political power of those with taxable wealth, and the level of natural resources. Rulers can attempt to extract revenues solely through coercive means, but evidence suggests that institutionalized bargaining and representative institutions are more conducive to robust revenue collection in the long run. The latter approach builds greater cooperation and compliance with tax policies.

In Tilly's (1990) framework, the presence of readily taxable resources is a critical factor that permits a capital-intensive extractive strategy driven by agreements with holders of capital. Robust long-distance trade, both internally and with external trading partners, is an engine of wealth creation provided an important source of revenues for the state, but excessive demands by the state could drive these merchants to different locations. Rulers instead found it worthwhile to negotiate with capital-holding classes in order to raise revenues for war (p. 64). In capital-poor areas, by contrast, strategies of coercive extraction prevailed.

This logic is consistent with the fiscal sociology perspective, which argues that variations in the source of revenue are strongly connected to the forms that states take. A central claim is that there are "strong synergies between (a) the degree of dependence of rulers on tax revenue, (b) the emergence of representative government, and (c) the strength and resilience of the state in the context of interstate competition" (Moore 2004, 299). Not only can these states borrow more readily when needs arise, but there is a governance dividend that emerges from the ongoing negotiations between states and citizens. Taxation leads to more accountable government and thus more effective states.

Following this logic, political systems that create deliberative, representative bodies in which different political factions have real power would be more conducive to the

construction of extractive capacities of the state. In addition to helping build political community, and thus enhance the legitimacy of the state, representative democracy creates linkages between rulers and citizens that facilitate the growth of state powers through consent. The result may be the kind of "compliance capacity" discussed in Berwick and Christia (2018).

The contrast is with rentier states that obtain revenues through other means. Revenues from oil, minerals, or international development assistance help satisfy the state's revenue imperative, but they do not require the kinds of state-society interactions that improve administrative capacities of the state. Empirically, then, we should observe that states with high dependence on natural resource extraction and foreign aid will have slower growth of state capacity than states that depend on taxation. Access to these revenues reduces the incentives to undertake the hard work of creating an administrative system capable of more sophisticated revenue collection methods, as well as reduces accountability pressures from citizens that would otherwise arise from raising revenues through taxation.

1.4 Summary

This section examines the connections between the core functions of states and the development of state capacity over time. It makes the argument that a holistic approach is necessary. To be effective and gain legitimacy, states must be able to provide security, create a strong political community, and satisfy the revenue imperative. Since these goals are linked together in many ways, a focus on either internal or external factors is incomplete. Rulers face incentives and constraints that lead them to implement policies that affect the development of state capacity, positively or negatively. The next step is to compile a set of empirical propositions that arise from this analysis to help sort through which factors matter most.

2 Some Testable Propositions

Given the scarcity of quantitative measures of state capacity, especially going back in time, most empirical work on state formation and capacity-building is qualitative in nature. I draw upon this fertile scholarship to specify propositions that can be tested with large-sample methods for the 1960-2015 period using the *State Capacity Dataset* (Hanson and Sigman 2013). The central goal of this analysis is to identify the factors that best explain changes in the levels of state capacity over time across countries.

A number of basic geographic and demographic characteristics are expected to facilitate or hinder the development of greater state capacity. These factors affect the ease with which states can penetrate their territories and collect revenues from populations. All else being equal, states should have more difficulty consolidating their hold over territories that are larger, that have fewer transportation linkages, and that have rougher terrain. Areas that are more distant from the center of political power are more likely to be disconnected from the political community and sit outside the easy reach of the state, particularly when transportation infrastructure is poor. Low population density, furthermore, raises the marginal costs of revenue collection.

Herbst (2014) explains the rationale for these propositions in his analysis of sub-Saharan Africa, where neither pre-colonial nor colonial states were constructed to control large territories. With low population density, land was relatively plentiful, so there was little incentive to build an infrastructure to defend particular borders. Although colonial powers created territorial states by drawing borders on maps, they often did very little to build an administrative network that extended beyond the capital cities, which were usually located on the coast. Transportation infrastructure, such as railways, were built for the purpose of resource extraction rather than to facilitate the creation of a coherent polity. Thus, newly-independent states were left with bureaucratic apparatuses that had little ability to penetrate throughout the territory. Where transportation linkages are more robust and population density is greater, however, states can more easily reach citizens.

Hypothesis 1 *Larger land area and rougher terrain hinder the growth of state capacity.*

Hypothesis 2 *Greater population density and access to waterways support the growth of state capacity.*

The construction of a political community characterized by vertical and horizontal legitimacy is hindered by factors that impede the formation of loyalty toward the state relative to local, religious, or kinship ties. These factors are likely to include high levels of social diversity (such as ethnic or linguistic diversity), status as a former colony, and the level of democracy. In the empirical literature, high levels of ethnic fractionalization are statistically correlated with poorer performance on a range of economic, political and human development indicators (Easterly and Levine 1997; Montalvo and Reynal-Querol 2005; Kimenyi 2006). Political competition along ethnic lines hinders the formation of attachments to the broader nation-state. Although state policies can affect ethnic identities over time (Weber 1976), the existing level of diversity is still relevant. Likewise, in countries where there are strong religious cleavages, and the rulers of the state are not able to draw upon religious unity as a source of legitimacy, state-building may be impeded.

Hypothesis 3 *Greater social diversity hinders the growth of state capacity.*

As discussed above, the challenges of creating a political community associated with the state are greater in former colonies, especially those whose borders do not align with pre-colonial institutions or the distribution of ethnic populations. Independence movements may not be sufficient, as the desire for independence from an external power is not the same thing as a national identity. Additionally, the institutional legacy of colonialism may be state structures developed to serve the interests of the colonizing power rather than govern a sovereign territory. On average, accordingly, we should expect that the growth of state capacity is slower in former colonies.

Hypothesis 4 *Post-colonial status is associated with slower growth of state capacity.*

Democratic governance in which citizens enjoy high levels of political rights, on the other hand, should help build the legitimacy of the state. As Moore (2004) contends, one effect of democracy is a governance dividend. Democratic states should build state capacity more quickly than non-democratic ones. Wang and Xu (2015), using the *State Capacity Dataset*, find that democratic political contestation is robustly associated with the growth of state capacity over time. This study will explore whether democracy remains robust once controlling for other factors.

Hypothesis 5 *In polities where levels of democracy and political rights are higher, state capacity should grow more quickly.*

Wars and rivalries threaten the security of the state. Rivalries, especially those that continue over a long period of time, should induce states to build coercive capacity and the administrative infrastructure to support national defense. The ongoing nature of the threats creates competitive pressure for continued accretion of capabilities, and these pressures should be stronger when states face multiple rivals. Rivalries focus political attention on security needs and the ability of the state to project power. Citizens will be more willing to support the allocation of resources toward building state power in these areas.

Hypothesis 6 *Growth of state capacity is faster for states that have enduring or multiple rivalries.*

The effect of wars could work in either direction, since the effect on state capacity likely hinges upon the outcome of the conflict. Provided the country does not face military defeat, wars with rivals likely have the same effects as rivalry in general. States also engage in isolated conflicts, however, where there is not an enduring rivalry. It is much less likely that these isolated conflicts would have the same positive effect on the growth of state capacity that rivalry has.

Hypothesis 7 *Isolated conflicts are not linked to subsequent changes in state capacity.*

Finally, as the analysis of the revenue imperative contends, states that depend upon natural resource rents or foreign assistance for revenue have fewer incentives to develop extensive taxation systems that typically lead to pressures for accountability and representative government. Although revenues gained from rents can support the construction of coercive capacities of the state, these states are less likely to benefit from the governance dividend predicted by the fiscal sociology literature. Administrative capacity is likely to grow more slowly.

Hypothesis 8 *Revenues from rents are associated with slower growth of administrative capacity.*

Each of these propositions has strong theoretical support from the literature on state formation and state capacity-building, though most of these accounts are focused on particular factors. With a more integrated empirical approach, we have the ability to determine whether some of the these factors are more important than others.

3 Methodology and Data

One advance in the present study, compared with previous work on this subject, is the use of the *State Capacity Dataset* from Hanson and Sigman (2013), updated to include annual estimates from 1960 to 2015 of the level of state capacity for all countries in the Polity dataset (Marshall and Jaggers 2009). These data were produced with a Bayesian latent variable analysis technique developed by Arel-Bundock and Mebane (2011) that uses Markov Chain Monte Carlo (MCMC) estimation methods. Specifically, a set of observed indicators believed to be strongly related to state capacity are employed to recover an estimate of this latent and unobserved concept.

In the update to Hanson and Sigman (2013), the authors use 22 different indicators represent three key dimensions of state capacity: administrative, extractive and coercive capacity. A full list of these indicators is provided in Table 1. Administrative

capacity is represented by a variety of measures: expert ratings of bureaucratic quality, census frequency, statistical capacity, and a revenue mixture that involves taxes on income rather than trade. Extractive capacities of the state are represented by the overall level of taxation, as well as ratings from the World Bank and other organizations. Finally, coercive capacities of the state are represented by military spending, regular military personnel, paramilitary and police personnel, ratings of law and order, and measures of state control over its territory.

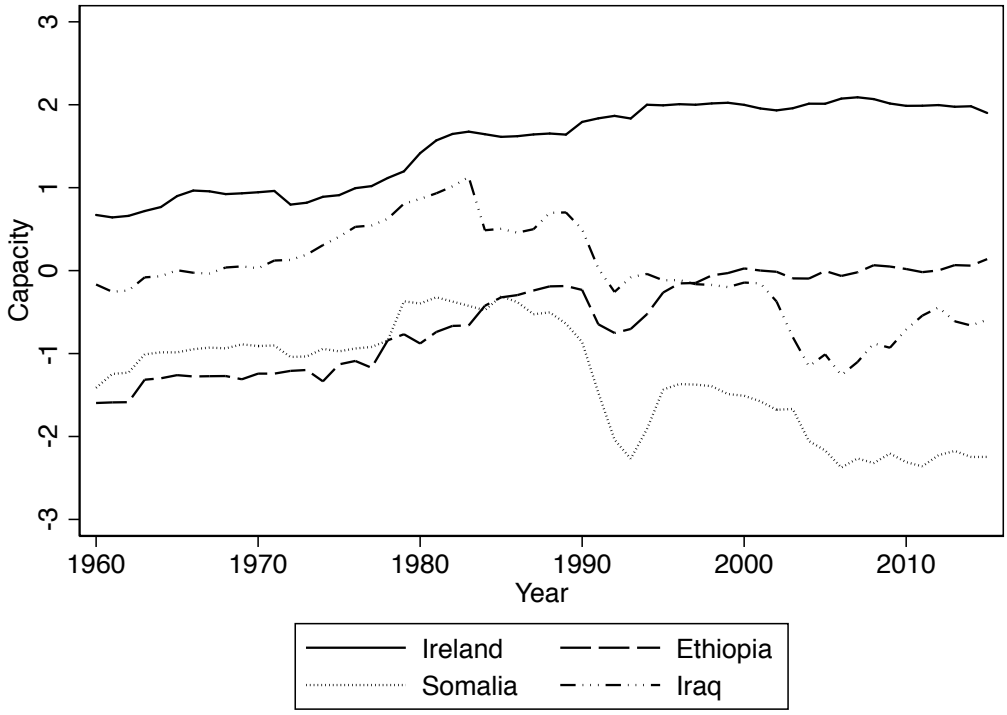
Table 1: Indicators of State Capacity

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 v8)	177	1960-2015
State Antiquity Index, based on Bockstette et al. (2002)	162	1960-2015
State Authority over Territory (V-Dem v8)	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
Weberianness (Rauch and Evans 2000)	34	1970-1990

The power of the MCMC method compared to traditional principal components analysis is its robustness to missing data. Rather than drop a country-year because one data point is missing, the method uses all available indicators for that country-year

to make the best estimate possible, along with the level of precision of this estimate. Thus, although data are more sparse in the early part of the 1960-2015 time period, we still can produce estimates of state capacity. The resulting dataset thus has far more extensive coverage than other efforts to measure bureaucratic quality or governance. *Capacity* is normalized to the standard normal scale, so a one-unit increase in the variables represents one standard deviation in the level of state capacity across countries.

Figure 1: Evolution of Capacity for Four Countries



There are a couple of limitations to keep in mind. First, the data are a bit noisy from one year to the next due to changing availability of the component indicators. Accordingly, it is better to examine longer-term trends in state capacity rather than annual changes. Second, the data are limited to the post-WW II era in which the world had been divided into territorial sovereign states with strong international norms against changing these borders. As such, the data are less well-suited to testing bellicist theories of state formation. We can, however, test whether rivalries and conflicts are asso-

ciated with changes in the level of state capacity.

As noted above, state-formation can take centuries. In the ideal, then, data would permit us to examine a much broader swath of time. It simply may be very difficult to observe the incremental changes in the level of state capacity that occur over a decade. With the *State Capacity Dataset*, we have the potential to examine changes in state capacity over a 55-year period for a moderately-sized sample of countries. Here, in order to obtain a larger sample that includes the post-colonial states that became independent during the 1960s, I instead look at the net change in Capacity 40-year period from 1975 to 2015.

Table 2: Largest Changes in Capacity 1975-2015

Largest Increases		Largest Declines	
Cape Verde Is.	+1.69	Nigeria	-0.30
Lesotho	+1.55	Rep. Congo	-0.34
Cambodia	+1.45	Central African Rep.	-0.52
Rwanda	+1.41	Guinea-Bissau	-0.60
Uganda	+1.40	Syria	-0.68
Bolivia	+1.39	Kuwait	-0.96
Uruguay	+1.30	Venezuela	-0.98
Ethiopia	+1.27	Iraq	-1.00
Spain	+1.24	Somalia	-1.27
Ecuador	+1.23	Libya	-1.82

Table 2 shows the ten countries with the largest increases in Capacity over this period, as well as the ten countries with the largest declines. Accordingly, there is significant variation in this measure across countries and over time. The empirical goal of this project is to explain this variation using variables that represent the theoretical perspectives outlined above.

I start with data on more enduring or slowly-changing country characteristics, such as geographical factors, drawn from a variety of sources. Descriptive statistics for these fixed variables are presented in Table 3.[1] Data on country land area (square km.),

[1]Of course, there are some changes in variables like ethnic fractionalization over time, but these changes tend to be slow. All these variables are treated as fixed throughout the time frame.

Table 3: Descriptive Statistics for Country Characteristics

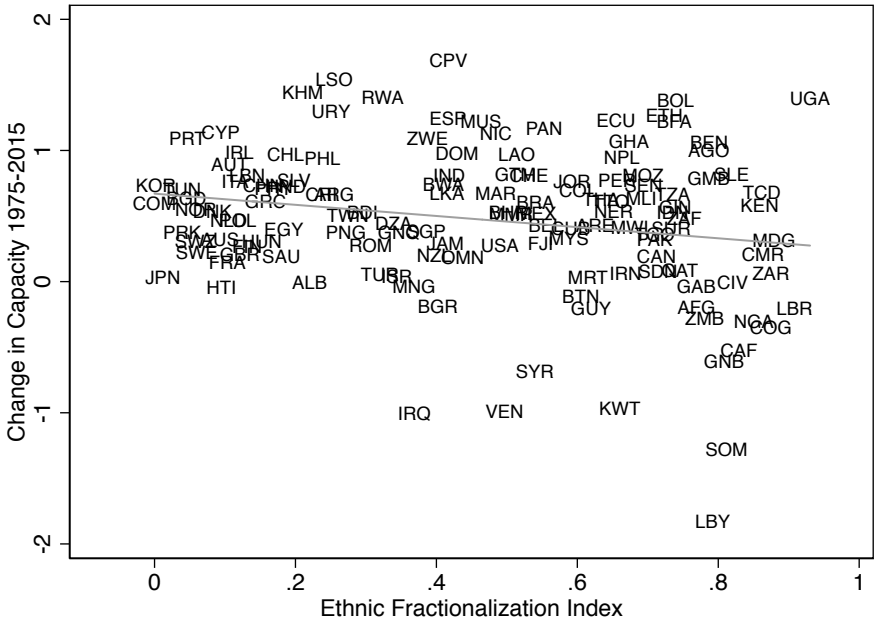
Variable	Obs	Mean	Std. Dev.	Min	Max
(log) Area	187	11.63	2.34	3.34	16.61
Colonized	177	0.82	0.38	0.00	1.00
Coastline	181	0.41	0.37	0.00	1.00
DistRiver	188	20.76	14.37	0.51	91.14
EthnicFrac	175	0.45	0.26	0.00	0.93
LangFrac	128	0.42	0.30	0.00	0.86
Rough	185	0.21	0.19	0.01	1.24
ReligPol	119	0.51	0.35	0.00	1.00
SoilSuit	167	0.56	0.21	0.00	0.96

soil quality (*SoilSuit*), and roughness of terrain (*Rough)* come from Ashraf and Galor (2013). SoilSuit is the proportion of the country's soil that is considered suitable for agriculture, and Rough measures terrain ruggedness (surface elevation changes) based upon geospatial data. The fraction of a country's borders that are coastal (*Coastline*) is calculated from public domain data. This captures access to sea transportation routes. Finally, the variable *DistRiver* is a geospatial measure of the mean distance to the nearest large river (Galor and Klemp 2015).

Another set of variables captures social characteristics that are invariant (as measured) during the study period. Three of these seek to capture different aspects of social diversity. The level of ethnic diversity (*EthnicFrac*) comes from Alesina et al. (2003). It is based on the usual Herfindahl index, although what constitutes an ethnic group is determined to some degree based upon country context. In some locales, language is a crucial differentiating factor, while race is more important in other contexts. Figure 2 plots the change in Capacity from 1975 to 2015 on EthnicFrac, showing a negative relationship between the two variables: Capacity tended to grow more slowly (or decline) in places that are more ethnically diverse.

The sample contains two other measures of social diversity. First, the variable *LangFrac*, which measures linguistic diversity, is also taken from Alesina et al. (2003). Second, a measure of societal polarization along religious lines (*ReligPol*) comes from

Figure 2: Change in Capacity and Ethnic Fractionalization



Montalvo and Reynal-Querol (2000). Unlike the measures of fractionalization, it gets larger when society becomes more evenly split into two religious groups.

The variable *Colonized* is a dichotomous measure of whether a country was ever colonized, drawn from the *ICOW Colonial History Dataset* (Hensel 2016). Since the analysis in this project includes independent countries, Colonized is meant to capture the effects of colonial legacy on the growth of state capacity post-independence. During the period 1975-2015, Capacity grew by 0.07 points slower on average in countries that had been colonized compared those that had not, but this difference is not statistically significant (p=.61).

A second set of variables measure country characteristics that can change more readily over time, at least in principle. Initially, these variables are measured annually. The annual-level data are then converted into periods of ten years in length. I thus seek to link changes in Capacity over ten years to country characteristics, whether during that same period or in the prior period.

The log value of population density (*PopDensity*), measured as persons per square

Table 4: Descriptive Statistics for Dataset with 10-year Time Periods

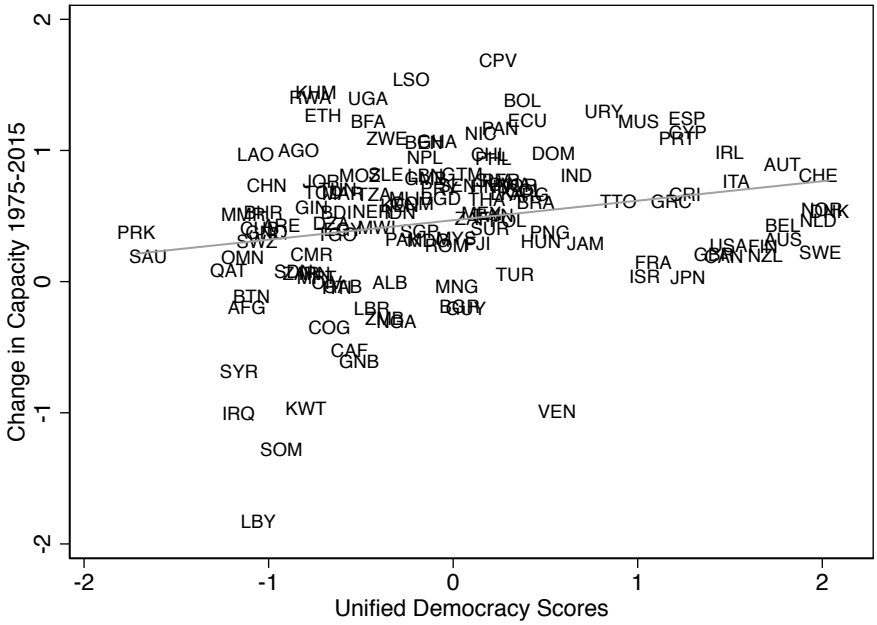
Variable	Obs	Mean	Std. Dev.	Min	Max
AllLeft	661	0.22	0.38	0.00	1.00
AllRight	665	0.09	0.25	0.00	1.00
AllCenter	666	0.03	0.14	0.00	1.00
Capacity	857	0.22	0.95	-2.31	2.82
ΔCapacity	682	0.16	0.35	-1.29	1.53
Conflicts	853	0.26	0.85	0.00	19.00
DemocUDS	921	-0.04	0.94	-2.02	2.08
GovFrac	618	0.18	0.24	0.00	1.00
ExecLeft	663	0.33	0.41	0.00	1.00
ExecRight	663	0.22	0.34	0.00	1.00
(log) FuelProduction	943	2.46	2.99	0.00	10.91
(log) GDPcap	903	8.72	1.25	5.78	12.38
Polarized	657	0.35	0.64	0.00	2.00
NATpct	827	4.56	7.31	-0.44	52.61
(log) PopDensity	895	3.80	1.57	-0.32	9.88
Rivalries	853	0.90	1.82	0.00	20.00

kilometer, comes from the World Development Indicators (World Bank 2016). When aggregating annual data to multi-year periods, I use the level of PopDensity at the beginning of the time period to avoid potentially endogenous relationships between the growth of state capacity and population density during that period.

There are many different measures available that relate to the concept of democracy. Most of them are problematic in one way or another. For each ten-year period, I use the mean value of the Unified Democracy Scores (*DemocUDS*) from Pemstein et al. (2010), a Bayesian latent measure of democracy that is constructed from 10 other commony-used indicators. In robustness checks (not presented), I find that the measure of political rights from Freedom House (2009) and Polity scores (Marshall and Jaggers 2009) produce very similar results. Figure 3 shows the bivariate relationship between the change in Capacity from 1975-2015 and the mean level of DemocUDS during this same period of time.

Other political variables code the left-right ideological position of the executive and ruling coalition using the *Database of Political Institutions* (Cruz et al. 2018). In the

Figure 3: Change in Capacity and Unified Democracy Scores



annual version of the data, the variables *ExecLeft* and *ExecRight* are indicator variables that equal 1 when the country's chief executive is coded as being of the political left or the political right respectively. The variables *AllLeft*, *AllRight*, and *AllCenter* are indicators coded as 1 in years when the executive and all houses of the legislature are controlled by parties with similar positions on the left-right spectrum. For the ten-year dataset, I use the means of these annual political indicators.

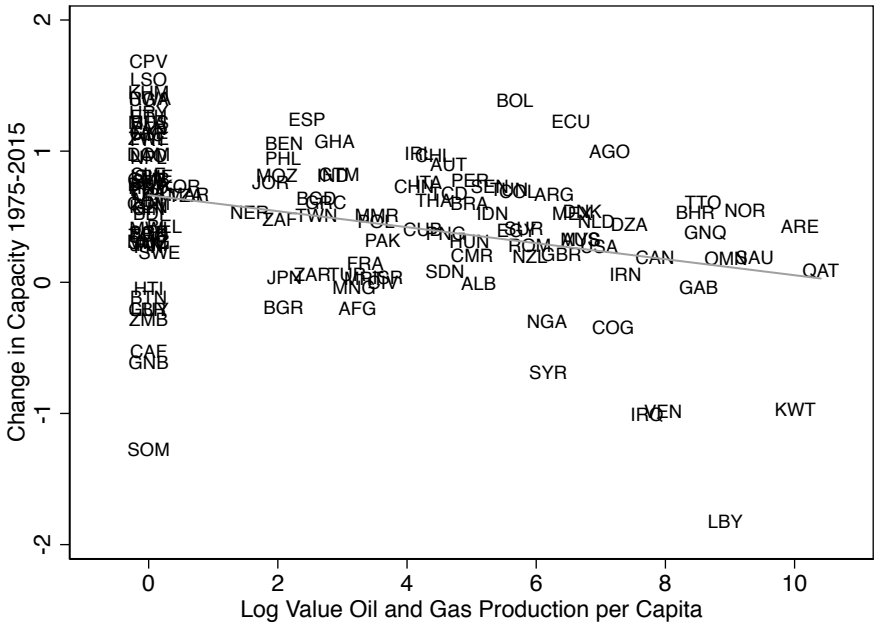
A third set of political variables measures political diversity in the government. The variable *GovFrac* measures the degree of party fractionalization in the governing coalition: it represents the probability that any two randomly-chosen legislators from the governing coalition will be of different parties. By contrast, the variable *Polarized* measures the distance between the party of the chief executive and the largest four parties in the legislature. For example, if the executive is of the political left and the largest legislative parties are of the political right, Polarized would be at its maximum of 2.

For data on international conflicts and rivalries, I turn to the New Rivalry Dataset

(Klein et al. 2006). In these data, hostile interactions between two countries are coded as either isolated conflicts or rivalries. For the variable *Conflicts*, I calculate the mean number of isolated conflicts with which a country is involved on an annual basis during each period of time. As argued above, this kind of conflict is less likely to be associated with growth in state capacity than an enduring rivalry. The variable *Rivalries* is the mean number of such rivalries had by each country during each period of time.

Ross and Mahdavi (2015) provide the underlying data for calculating the variable *FuelProduction*, which is the mean of the log value of oil and gas production per capita over the years in each time period. Since the data are measured on a per capita basis, the variable provides a relative comparison of the impact of oil and gas in the country's overall economy. In Figure 4, we see the relationship between the change in Capacity from 1975-2015 and the mean level of FuelProduction during this time period.

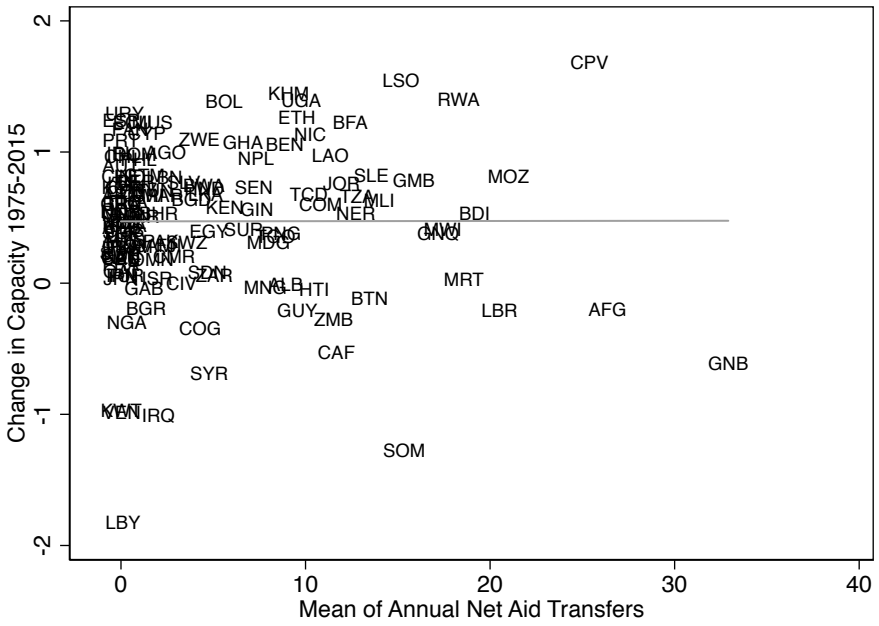
Figure 4: Change in Capacity and Fuel Production per capita



Data on the level of international assistance comes from Roodman (2005). The variable *NATpct* is net aid transfers, including loan cancellation, measured as a percentage of GDP. Donor countries are coded as having a value of 0 on this variable. The larger

the amount of assistance as a percentage of the economy, the more aid-dependent is the country. As Figure 5 shows, there is not a strong relationship between the change in Capacity from 1975-2015 and the mean level of net aid transfers during this time period.

Figure 5: Change in Capacity and Net Aid Transfers



Finally, data on GDP per capita come from the Penn World Tables version 9.0 (Feenstra et al. 2015). I use the log value of real GDP per capita calculated using national accounts (rgdpna). The lagged value of this variable may help serve to control for prior success at state building, allowing for better identification of the effect of the other independent variables on the subsequent change in the level of state capacity.

4 Multivariate Analysis with Ten-Year Periods

In each set of multivariate tests, the dependent variable is ΔCapacity, the change in the level of state capacity during each ten-year period. I start with a baseline model in which the key independent variables are basic country characteristics: geographical

factors and population density. Each model also controls for the level of Capacity$_{t-1}$ and the log value of GDP per capita at the end of the previous period. To this baseline, I then add sets of independent variables that are associated with the different theoretical perspectives discussed above. The final set of models is more comprehensive in nature.

4.1 Baseline Model and Social Diversity Measures

Table 5, Model 1 shows the estimated coefficients for the baseline model. To facilitate interpretation, it is useful to note that the mean of ΔCapacity is .16, indicating that Capacity on average grew by .16 points over each ten-year period. In general, then, we can interpret the coefficients on all independent variables as adjustments to the rate of change in Capacity over time. The two key variables that control for the baseline condition of the country prior to each time period – Capacity$_{t-1}$ and lnGDPcap$_{t-1}$ – are statistically and substantively significant in all models. Where country wealth as measured by log GDP per capita is high, Capacity grows faster in the subsequent period. On the other hand, the rate of change in Capacity is slower where Capacity is already high, holding all other factors constant. Combined, these coefficients tell a coherent story.

To illustrate, suppose a country is at 0 on Capacity and has $3,000 in GDP per capita (lnGDPcap = 8.01). All else equal, the predicted effect of lnGDPcap on ΔCapacity is to increase it by about .48. If Capacity were at 1, however, this predicted rise in ΔCapacity would be tempered by a .15 reduction due to its higher pre-existing level of Capacity. Since bivariate correlation between Capacity and lnGDPcap is quite high (r=0.78), these two variables tend to cancel each other out. A rise in GDP per capita is associated with a positive shift in ΔCapacity, but the effect tapers off once the level of Capacity increases.

Of population density and the five geographic factors, only the mean distance to the nearest large river was found to have a statistically significant with respect to ΔCapacity. An increase in DistRiver of one standard deviation (14.37) is associated

Table 5: Base Model and Social Diversity Measures

	(1)	(2)	(3)	(4)	(5)
PopDensity	0.01	0.00	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
LandArea	0.00	0.01	0.00	0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SoilSuit	0.12	0.04	0.15*	0.11	0.05
	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
Rough	0.10	0.09	0.11	0.12	−0.00
	(0.10)	(0.11)	(0.10)	(0.10)	(0.11)
Coastline	0.05	0.04	0.06	0.05	0.07
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
DistRiver	0.00**	0.00**	0.00**	0.00**	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
EthnicFrac		−0.16**			
		(0.06)			
LangFrac			−0.05		
			(0.04)		
ReligFrac				0.07	
				(0.04)	
ReligPol					−0.17**
					(0.05)
Capacity_{t-1}	−0.15**	−0.17**	−0.16**	−0.15**	−0.16**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
lnGDPcap_{t-1}	0.06**	0.06**	0.07**	0.07**	0.05*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	−0.57*	−0.44^	−0.58*	−0.61**	−0.21
	(0.23)	(0.24)	(0.24)	(0.23)	(0.26)
N	565	563	551	565	496
Countries	142	141	139	142	108
R^2	0.08	0.09	0.09	0.08	0.09

Table 5. Random-effects OLS model with panel-clustered standard errors. Each observation is a ten-year period for each country. The dependent variable is Δ Capacity_t: the change in Capacity from the previous period to the current period.

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

with a small increase in Capacity by .036 over each ten-year period. This result is contrary to theoretical expectations that waterway access facilitates state penetration of the territory, but it appears that the finding is a statistical artifact. It is not robust to the addition of other variables in subsequent sets of tests.

Models 2 through 5 on Table 5 sequentially test four variables measuring social conditions that may affect the establishment of a stronger sense of political community associated with the state. Two variables expected to have a negative effect on the construction of state capacity – EthnicFrac and ReligPol – indeed have negative and statistically significant coefficients, while LangFrac and ReligFrac do not. The findings for the religion variables have certain logic: religious diversity does not affect the growth of state capacity, but religious polarization does. Where the level of religious polarization as measured by ReligPol is one standard deviation (.35) higher, the predicted level of ΔCapacity is about .06 points lower over ten years. Similarly, where EthnicFrac is one standard deviation (.26) higher, the change in Capacity is expected to shift in the negative direction by about .04 over ten years, all else being equal.

The coefficient on LangFrac is negative, which is consistent with theoretical expectations, but its standard error is similar in magnitude. One interpretation of this result is that ethnic differences, including race-based differences, have a more powerful (negative) effect on the creation of a sense of shared political community under the state than do linguistic differences.

4.2 Political Factors

Table 6 presents results from models that incorporate measures of various political characteristics. Each of these five models continues to include the log of population density and the various measures of geographical factors, but the coefficients are not reported in order to save space.[2] Models 2 through 5 examine different sets of variables

[2]None of these coefficients were statistically significant at conventional levels of significance.

Table 6: Politcal Factors

	(1)	(2)	(3)	(4)	(5)
Colonized	−0.10**	−0.08*	−0.08*	−0.08*	−0.08*
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)
DemocUDS	0.09**	0.09**	0.09**	0.10**	0.07*
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
ExecLeft		0.02			
		(0.04)			
ExecCenter		0.04			
		(0.08)			
AllLeft			0.01		
			(0.04)		
AllRight			0.01		
			(0.06)		
AllCenter			−0.15		
			(0.14)		
GovFrac				0.12^	
				(0.06)	
Polarized					0.05^
					(0.03)
Capacity_{t-1}	−0.21**	−0.23**	−0.23**	−0.22**	−0.22**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
lnGDPcap_{t-1}	0.06**	0.07**	0.07**	0.06**	0.06**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	−0.18	−0.35	−0.43	−0.43	−0.33
	(0.25)	(0.27)	(0.28)	(0.27)	(0.26)
N	563	482	481	450	479
Countries	141	140	140	141	139
R^2	0.12	0.13	0.13	0.15	0.13

Table 6. Random-effects OLS model with panel-clustered standard errors. Each observation is a ten-year period for each country. The dependent variable is Δ Capacity_t: the change in Capacity from the previous period to the current period. Controls for population density and geographical factors from Table 5 are included but coefficients not reported here.

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

that measure left vs. right control of government and the degree of political party fractionalization or polarization. The variables Colonized and DemocUDS are included in all five models.

As a reminder, Colonized is a dichotomous indicator of whether the country is a former colony. The key finding is clear and consistent across all models: ΔCapacity is lower in former colonies by about .08 to .10 points on average. Over time, these effects accumulate and lead to substantively significant differences in state capacity. These findings are consistent with the account of Herbst (2014), who argues that colonial states in many cases were designed to serve the interests of the colonizing power and suffered from legitimacy. This legacy continues to undermine state-building efforts post-independence.

The effect of democracy, on the other hand, consistently works in the other direction. For each increase in DemocUDS by one standard deviation, the predicted value of ΔCapacity rises by .07 to .10 points according to the five models presented here. Robustness checks using several other democracy-related measures, such as Polity or Freedom House scores, were similarly consistent. State capacity appears to grow faster in countries that exhibit more characteristics of democracy, confirming the findings of Wang and Xu (2015). This is consistent with theories that connect representative democracy to the construction of political community and with the fiscal sociology perspective.

In order to test theories that connect the construction of state capacity with the ideological characteristics of those in control of the state (e.g. Grassi and Memoli, 2016), Model 2 includes the mean level of the variables ExecLeft and Exec Right over each ten-year period. Both coefficients are smaller than their standard errors, indicating the inability to identify a systematic relationship. In Model 3, I look beyond the executive branch and code scenarios of unified control over the government by parties of the left, right or political center. None of these variables appear to be strongly connected to the change in Capacity either.

Rather than unified control, it appears that dividing control across different political factions is connected with positive changes in ΔCapacity. As Model 4 of Table 6 shows, the coefficient on GovFrac, which measures party fractionalization in the governing coalition, is positive and significant at the .1 level. Where GovFrac is one standard deviation (.24) higher, ΔCapacity is expected to be about .03 points higher. Similarly, greater ideological distance between the executive and the largest parties in the legislature, as measured by Polarized, is associated with faster growth of Capacity.

In sum, political systems that create democratic political competition and bring multiple perspectives into government appear be more conducive to the construction of state capacity than are their non-democratic counterparts. This effect does not appear to be tied to the ideology of either the chief executive or the ruling coalition.

4.3 Rivalry and Conflict

Table 7 brings a test of the bellicist perspective, adding the variables Conflicts and Rivalries. As argued above, rivalries are more enduring and are expected to drive faster growth of Capacity, whereas isolated conflicts are less likely to produce ongoing incentives to build state capacity and can be harmful. Neither variable is significant when used in isolation, but both are significant when included together (Model 4). The results are consistent with these expectations.

Each additional rivalry is associated with a ten-year change in Capacity that is .02 points greater, while isolated conflicts have a negative effect on Capacity. In robustness checks not reported here, conflicts in the more distant past continue to be associated with slower growth in Capacity, though the estimates are not statistically significant. Nevertheless, the positive effect Rivalries provides evidence that, even though the post-WW II era is one in which the international system has largely set country borders, steady security concerns still lead to construction of state capacity.

Table 7: Measures of Conflict and Rivalry

	(1)	(2)	(3)	(4)
Conflicts	−0.05			−0.07^
	(0.04)			(0.04)
Conflicts_{t-1}		−0.03		
		(0.03)		
Rivalries			0.01	0.02*
			(0.01)	(0.01)
Capacity_{t-1}	−0.15**	−0.15**	−0.15**	−0.15**
	(0.02)	(0.02)	(0.02)	(0.02)
lnGDPcap_{t-1}	0.07**	0.07**	0.07**	0.07**
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	−0.64**	−0.64**	−0.57*	−0.59*
	(0.23)	(0.23)	(0.24)	(0.23)
N	560	561	560	560
Countries	141	141	141	141
R^2	0.08	0.08	0.08	0.08

Table 7. Random-effects OLS model with panel-clustered standard errors. Each observation is a ten-year period for each country. The dependent variable is Δ Capacity_t: the change in Capacity from the previous period to the current period. Controls for population density and geographical factors from Table 5 are included but coefficients not reported here.

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

4.4 Rents from Resource Extraction and Aid Revenues

In Table 8, I present tests that examine the effects of access to resource-related revenues and flows of international assistance. Model 1 shows a modest, negative relationship between the change in Capacity and the (log) per capita value of oil and gas production in a country during the time period. A one-standard deviation (2.99) increase in FuelProduction is associated with a .04 reduction in ΔCapacity. By contrast, resources coming from diamond production have a modest positive association with change in Capacity. For DiamondProd, a one-standard deviation (0.65) increase leads to a predicted rise in ΔCapacity by about .01. It is not easy reconcile these two findings, especially given the association between diamonds and conflict. On the other hand, diamond production is concentrated in a small number of countries, and the results appear to be driven heavily by Botswana.

International assistance during a ten-year period is associated with slower growth (or potentially a decline) of Capacity. Take two countries that are identical in all respects except that the first country receives no net foreign assistance, while level of foreign assistance to the second country equals ten percent of GDP. Capacity is expected to be about .1 points lower in the second country after ten years. The direction of the causal arrows when it comes to foreign assistance is not entirely clear, however. Logically, foreign assistance levels would rise in response to a significant crisis that the recipient state cannot manage. Rather than cause a decline of state capacity, in other words, the aid is a response to state weakness or failure at some level. A scenario of this kind could produce a statistical association identical to what is observed in the data. When using lagged values of NATpct, the statistical significance of the coefficient disappears.

4.5 Comprehensive Models

Table 9 brings together the most-significant variables from the previous results to present four comprehensive models that represent the different theoretical perspec-

Table 8: Measures of Extraction and Aid Revenues

	(1)	(2)	(3)
FuelProduction	−0.01**		
	(0.01)		
DiamondProd		0.02**	
		(0.01)	
NATpct			−0.01^
			(0.00)
Capacity_{t-1}	−0.16**	−0.15**	−0.14**
	(0.02)	(0.03)	(0.02)
lnGDPcap_{t-1}	0.09**	0.07*	0.04^
	(0.02)	(0.03)	(0.02)
Constant	−0.82**	−0.57*	−0.19
	(0.24)	(0.27)	(0.29)
N	562	415	536
Countries	142	118	141
R^2	0.08	0.05	0.08

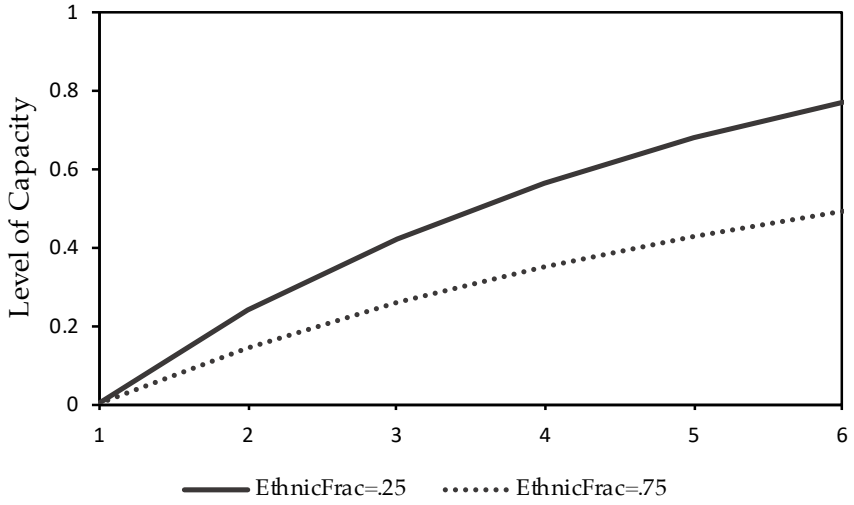
Table 8. Random-effects OLS model with panel-clustered standard errors. Each observation is a ten-year period for each country. The dependent variable is Δ Capacity_t: the change in Capacity from the previous period to the current period. Controls for population density and geographical factors from Table 5 are included but coefficients not reported here.

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

tives outlined above. Overall, the results support the conclusion that growth in state capacity arises from a variety of factors. Rather than discuss each model sequentially, I will summarize the key findings from across them.

First, among the factors explored in this analysis, post-colonial status and ethnic diversity are the key challenges to the growth of state capacity. These variables remain statistically significant even after the inclusion of other factors. Both of these variables are related to the creation of a political community under the state, showing that underlying conditions can impede state-building efforts. Religious polarization does not appear nearly as important, either in terms of magnitude or statistical significance, one we have accounted for other factors. Figure 6 uses the results from Model 4 on Table 9 to illustrate the predicted effect of EthnicFrac for a hypothetical scenario in which the other variables are set to their median values and the log value of GDP per capita grows at the median rate.

Figure 6: Predicted Evolution of Capacity at Different Levels of EthnicFrac



Second, democracy is consistently and robustly associated with more rapid growth in state capacity. In particular, democracies in which there is diversity in the governing party coalition are the most conducive to rapid growth in state capacity. These conditions are met in places like Belgium (late 1980s-2000s), Brazil (1990-2010), Fin-

Table 9: Comprehensive Models

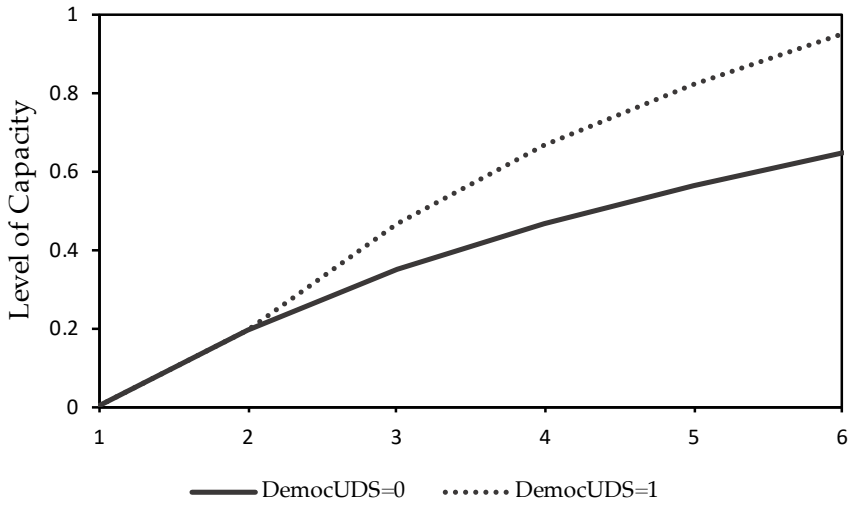
	(1)	(2)	(3)	(4)
Colonized	−0.07*	−0.07*	−0.05^	−0.05^
	(0.03)	(0.03)	(0.03)	(0.03)
DemocUDS	0.13**	0.15**	0.13**	0.12**
	(0.02)	(0.02)	(0.03)	(0.03)
EthnicFrac	−0.18**	−0.21**	−0.12*	−0.19**
	(0.06)	(0.08)	(0.06)	(0.06)
ReligPol		−0.05		
		(0.05)		
GovFrac			0.14*	
			(0.06)	
Polarized				0.04
				(0.03)
Conflicts	−0.00	0.02	−0.00	−0.01
	(0.05)	(0.05)	(0.05)	(0.05)
Rivalries	0.03**	0.03**	0.02^	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
FuelProduction	−0.01	−0.01	−0.01	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)
NATpct	−0.00	−0.00	−0.00	−0.01
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.09	0.20	−0.02	0.14
	(0.32)	(0.35)	(0.33)	(0.35)
N	528	471	426	451
Countries	139	107	138	137
R^2	0.18	0.19	0.20	0.18

Table 9. Random-effects OLS model with panel-clustered standard errors. The dependent variable is Δ Capacity$_t$: the change in Capacity from the previous period to the current period. Controls for the log level of GDP per capita. Controls for population density and geographical factors from Table 5 are included but coefficients not reported here.

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

land (1970-2010), and Switzerland (1970-2010). In these results, ideological distance (i.e. polarization) between the executive and the legislature is no longer related to the changes in Capacity with high statistical confidence. The effect over time of increasing DemocUDS from 0 to 1 is illustrated in Figure 7.[3]

Figure 7: Effect of Increasing DemocUDS by 1 at t=2



Third, enduring international rivalries do have a mild effect in promoting the growth of Capacity. Actual conflicts do not appear to have much of an impact at all. Most likely, rivalries are associated with higher levels of military spending and greater troop numbers, both of which are indicators of coercive capabilities in the Capacity measure. Additionally, one can speculate that the insecurity created by rivalries can produce political focus and incentives to make improvements in state capabilities over time. With respect to conflicts, the measure may simply be too blunt, since it codes neither for where the conflict was fought nor its outcome.

Finally, the variables representing access to natural resource rents or international assistance do not appear to be influential on the change in state capacity over time. Once we control for other critical factors, these coefficients attenuate to zero and lose statistical significance. These findings cast doubt upon the fiscal sociology perspective.

[3]Uses estimates from Table 9, Model 4. All other variables set to their median values, and the log value of GDP per capita grows at the median rate.

Overall, then, the evidence is consistent with the argument that a range of factors affect the incentives for building state capacity and the ease with which capacity can be increased. More work on bringing together the somewhat disparate theoretical perspectives is needed.

5 Conclusion

This paper argues for a holistic approach toward understanding what factors affect the growth of state capacity over time. The vibrant literature on this topic tends to focus on particular pieces of the puzzle: the role of conflict in the international system, the need for revenue, and domestic factors that affect the relationship between citizens and the state. Since they are not mutually exclusive, there is value in bringing them together into a single framework.

Empirically, the approach taken in this paper thus gives equal treatment to the variety of factors that may affect state-building efforts, using large-sample methods to determine which factors appear to matter the most. The results may offer some guidance for developing more integrated theories. As the findings from the comprehensive models indicate, more attention should be devoted to understanding the interrelationships between regime characteristics and underlying social conditions. There is strong evidence that regime characteristics are important. Most likely, regimes that provide more inclusive political representation are needed to help build political community and facilitate compliance with the state in the face of diverse social circumstances and unfavorable historical-institutional legacies.

Additionally, in the contemporary era, some rethinking of the bellicist perspective may be necessary. Although rivalries and security challenges continue to play a role in creating incentives for building state capabilities, it does not appear that "conflict" can be generally linked to a clear effect on state capacity. Future work can disaggregate types of conflict, coding for the winners/losers and noting whether the conflicts are about defining territorial boundaries.

A comprehensive theory is not articulated in this paper, but minor modifications to existing theories will suffice to create a framework to move this research forward. As set forth in Levi (1988), we can think of rulers as rational actors that seek to maintain their power and maximize their revenues. They face threats of many kinds, some from rivals, some from society, and some from external sources. They also face constraints that are created by institutions, geography, the nature of society, and so forth.

In response, rulers produce policies that lead to changes in state capacity over time. They may seek to build the state's military capabilities, extend the administrative infrastructure to deliver new kinds of public services, or build new revenue-collection mechanisms. Due to constraints, however, these policies may not be workable, or the need to cater to powerful constituents may lead to deleterious effects on state capacity over the long run.

References

Adelman, Irma and Cynthia Taft Morris. 1967. *Society, Politics, & Economic Development: A Quantitative Approach*. Baltimore: The Johns Hopkins Press.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2): 155–194.

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.

Ashraf, Quanmrul and Oded Galor. 2013. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Development." *American Economic Review* 103(1): 1–46.

Berwick, Elissa and Fontini Christia. 2018. "State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate." *Annual Review of Political Science* 21: 71–91.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth* 7: 347–369.

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2016. "The Lay of the Land: Information Capacity and the Modern State." (STANCE Working Paper Series; Vol. 2016, No. 2). Lund: Department of Political Science: Lund University.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, , and Daniel Ziblatt. 2018. "V-Dem Codebook v8." URL `https://doi.org/10.23696/vdemcy18`. Varieties of Democracy (V-Dem) Project.

Cruz, Cesi, Philip Keefer, and Carlos Scartascini. 2018. *Database of Political Institutions 2017 (DPI2017)*. Inter-American Development Bank. Numbers for Development. URL `https://mydata.iadb.org/Reform-Modernization-of-the-State/ Database-of-Political-Institutions-2017/938i-s2bw`.

Dannreuther, Roland. 2007. "War and Insecurity: Legacies of Northern and Southern State Formation." *Review of International Studies* 33(2): 307–326.

Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *The Quarterly Journal of Economics* 112(4): 1203–1250.

Ertman, Thomas. 1997. *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge: Cambridge University Press.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10): 3150–3182. URL `www.ggdc.net/pwt`.

Freedom House. 2009. *Freedom in the World Country Ratings: 1972 to 2008*. Washington, D.C.: Freedom House, Inc.

Galor, Oded and Marc Klemp. 2015. "Roots of Autocracy." Unpublished Manuscript.

Grassi, Davide and Vincenzo Memoli. 2016. "Political Determinants of State Capacity in Latin America." *World Development* 88: 94–106.

Hanson, Jonathan K. and Rachel Sigman. 2013. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." Unpublished Manuscript.

Hensel, Paul R. 2016. "ICOW Colonial History Dataset, version 1.0." Available at <http://www.paulhensel.org/icowcol.html>.

Herbst, Jeffrey. 2014. *States and Power in Africa*. Princeton: Princeton University Press.

Hintze, Otto. 1975a. "Formation of States and Constitutional Development." In Felix Gilbert, ed., *The Historical Essays of Otto Hintze*. Oxford: Oxford University Press.

———. 1975b. "Military Organization and State Organization." In Felix Gilbert, ed., *The Historical Essays of Otto Hintze*. Oxford: Oxford University Press.

Holsti, Kalevi J. 1996. *The State, War, and the State of War*. Cambridge: Cambridge University Press.

Kimenyi, Mwangi S. 2006. "Ethnicity, Governance and the Provision of Public goods." *Journal of African Economies* 15: 62–99.

Klein, James P., Gary Goertz, and Paul F. Diehl. 2006. "The New Rivalry Dataset: Procedures and Patterns." *Journal of Peace Research* 43(3): 331–348.

Levi, Margaret. 1988. *Of Rule and Revenue*. Berkeley and Los Angeles: University of California Press.

Marshall, Monty G. and Keith Jaggers. 2009. *Polity IV Project: Dataset Users' Manual*. Center for Systemic Peace, George Mason University.

Montalvo, Jose G. and Marta Reynal-Querol. 2000. "The Effect of Ethnic and Religious Conflict on Growth." Unpublished Manuscript. Department of Economics, Universitat Pompeu Fabra.

———. 2005. "Ethnic diversity and economic development." *Journal of Development Economics* 76(2): 293–323.

Moore, Mick. 2004. "Revenues, State Formation, and the Quality of Governance in Developing Countries." *International Political Science Review* 25(3): 297–319.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Pemstein, Daniel, Stephen A. Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4): 426–449.

Rauch, James and Peter Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75(1): 49–71.

Roodman, David. 2005. "An Index of Donor Performance." Working Paper 67, Center for Global Development.

Ross, Michael and Paasha Mahdavi. 2015. "Oil and Gas Data, 1932-2014." doi: doi/10.7910/DVN/ZTPW0Y. URL http://dx.doi.org/10.7910/DVN/ZTPW0Y.

Spruyt, Hendrik. 1994. *The Sovereign State and its Competitors: An Analysis of Systems Change*. Princeton: Princeton University Press.

———. 2011. "War, Trade, and State Formation." In Robert E. Goodin, ed., *The Oxford Handbook of Political Science*. Oxford: Oxford University Press.

Strayer, Joseph. 1970. *On the Medieval Origins of the Modern State*. Princeton: Princeton University Press.

Thomson, Janice E. 1994. *Mercenaries, Pirates, and Sovereigns: State-Building and Extraterritorial Violence in Early Modern Europe*. Princeton: Princeton University Press.

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990-1990*. Cambridge: Blackwell.

Wang, Erik H. and Yiqing Xu. 2015. "Awakening Leviathan: the Effect of Democracy on State Capacity, 1960-2009." Unpublished Manuscript.

Weber, Eugen. 1976. *Peasants into Frenchmen: The Modernization of Rural France 1870-1914*. Stanford: Stanford University Press.

World Bank. 2016. *World Development Indicators 2016 (WDI Online)*. Washington, D.C.: The World Bank.