

Instructions for authors, subscriptions and further details:

http://rimcis.hipatiapress.com

State Capacity: Utilization, Durability, and the Role of Wealth vs. History

Elaine Enriquez & Miguel Angel Centeno¹

1) Department of Sociology, Princeton University, United States

Date of publication: November 30th, 2012

To cite this article: Enriquez, E., & Centeno, M.A. (2012). State Capacity: Utilization, Durability, and the Role of Wealth vs. History. *International and Multidisciplinary Journal of Social Sciences, 1*(2), pp. 130-162, doi: 10.4471/rimcis.2012.07

To link this article: http://dx.doi.org/10.4471/rimcis.2012.07

PLEASE SCROLL DOWN FOR ARTICLE

The terms and conditions of use are related to the Open Journal System and to Creative Commons Non-Commercial and Non-Derivative License. *RIMCIS – International and Multidisciplinary Journal of Social Sciences Vol. 1 No. 2 November 2012 pp. 130-162*

State Capacity: Utilization, Durability, and the Role of Wealth vs. History

Elaine Enriquez & Miguel Angel Centeno Princeton University

Abstract

The concept of state capacity has been in development literature for decades. Nevertheless the concept, its operationalization, and its measurement are still highly contested. This paper seeks to briefly review the literature on state capacity and provide an empirical analysis of recent data in order to re-assess the state of capacity theory and testing. We argue that very little, if any, attention has been paid to critical variations in national regional and subnational levels in state service provision, both statically and over time. We also argue that existing theoretical research in capacity utilization can provide insights to state building and development scholars regarding who is "doing more with less". Finally, we offer the concept of durable capacity as a way of understanding how states can provide long-standing development outcomes despite income variations.

Keywords: state capacity, state power, state strength, state building

2012 Hipatia Press ISSN 2014-3680 DOI: 10.4471/rimcis.2012.07



RIMCIS – International and Multidisciplinary Journal of Social Sciences Vol. 1 No. 2 November 2012 pp. 130-162

Capacidad del estado: Utilización, duración y rol de la riqueza vs. historia

Elaine Enriquez & Miguel Angel Centeno Princeton University

Resumen

El concepto de capacidad del estado ha sido desarrollado en la literatura durante décadas. A pesar de esto, el concepto, su operacionalización y su medida aún son muy discutidos. Este artículo revisa el concepto de capacidad del estado y proporciona un análisis empírico de datos recientes a fin de reformular la teorización y la medida de la capacidad del estado. En el mismo argumentamos que se ha prestado muy poca atención, si es que se ha prestado alguna, a los puntos de inflexión críticos en los cambios nacionales, regionales y subnacionales en cuanto a provisión de servicios por parte del estado, tanto estadísticamente como con el paso del tiempo. Este trabajo es el primero en examinar a nivel nacional y subnacional, la capacidad como una variación en la provisión de servicios.

Palabras clave: capacidad del estado, poder del estado, fortaleza del estado, construcción del estado

2012 Hipatia Press ISSN 2014-3680 DOI: 10.4471/rimcis.2012.07



• he concept of state capacity has existed in development literature for decades. Development practitioners have declared capacity development a fundamental project. It served as a key topic at the High Level Forums held by the OECD in Paris in 2006 and Accra in 2008 and the G8 summit in Geneagles in 2005. As stated by the OECD (2006): "...capacity development is a fundamental component of development and aid effectiveness and a key element in achieving the Millennium Development Goals (MDGs)" (p. 3). Further, "Adequate country capacity is one of the critical missing factors in current efforts to meet the Millennium Development Goals (MGDs)" (p. 7). Academics in politics, comparative history, and sociology have acknowledged the importance of understanding capacity as well. The literatures on state capacity and its related concepts are vast. We can point to two recent publications, a special issue of Studies in Comparative International Development dedicated to infrastructural power (for the editorial introduction, see Soifer & Hau, 2008), and the extended exchange between the developers of the World Bank's World Development Indicators and its critics in the Journal of Politics (see Kaufmann, Kraay, & Mastruzzi, 2007a, 2007b; Kurtz & Schrank, 2007a, 2007b) as indicators of the importance of the concept and its measurement to the academic community.

Surprisingly, there has been limited work on developing a coherent and empirically verifiable concept that can receive broad support. Even after significant time and empirical research there remain competing definitions, competing hypotheses, and competing methods of measurement. The positive outcome of this healthy debate is the assurance that the concepts of state capacity and state strength remain an important and critical to academics in comparative politics and practitioners of development. However, without further theoretical and empirical development, capacity will cease to be a productive concept (for an argument that the capacity concept has already ceased to be productive, see Kocher, 2010).

The role of the state in development outcomes is widely accepted (Evans, 1995; Kohli, 2004; Przeworski, 1990). The coordination and resources needed to enact widespread improvements in the quality of life and sustained economic growth have been widely viewed as best

met through the role of the developmental state. While there are arguments for the role of free trade to promote economic development, the inability of free markets to limit concomitant inequality with growth and the limits of free markets to provide social development have brought us back to the need of the state to be a key agent of development (Collier, 2007; Stiglitz, 2003). Nevertheless, the state is a political entity and thus has interests of its own separate from the interests of the citizens over which it presides. Further, the state must cope with conflicts inherent in a complex organization that seeks to maintain itself over time. This says nothing of the changing character of the state and expectations that citizens have of the role of state and what it should and should not do.

The concept of state capacity is tied to the relatively modern understanding that the state is, at least in part if not fundamentally, responsible for the well-being, economic status, and social development of its citizens. This is to say that a modern citizen now expects more from its state that citizens of times past. While not analyzed extensively in this paper, what this means for the theoretical understanding of state capacity is that we must ask the questions: State capacity to do what and for whom? To answer the first question, this paper is based on the assumption that the states analyzed are interested in providing basic services to their people as defined by the United Nations' Millennium Development Goals (UN, 2000). The data analyzed in this paper do not allow us to answer the second question of for whom? Nevertheless, we would argue that in most cases states provide services to the population at differential rates, and those differences can be detected and related to a host of other factors. We will discuss these questions at the conclusion of the paper.

This paper seeks to make a comprehensive analysis and evaluation of the concept of capacity, highlighting its varied use in the literature both conceptually and empirically. After reviewing the theoretical variations of the concept, we use an empirical analysis to highlight conceptual and measurement problems of capacity. Our theoretical and empirical review leads us to believe that the future of empirical analysis of the developmental state will require case-based research that accounts for regional and historical contextualization. Particular attention must be paid to theoretical definitions and empirical operationalizations that differentiate capacity vs. utilization and differentiate the inputs from the outputs.

What is Capacity?

Political and sociological literature discussing the state regularly uses the concept "state capacity" and related terminology and ideas, such as "power," "state strength," and "institutions." The notion of state capacity in its various forms has existed for decades and was an element of much of 19th and 20th century German social theory, but it became a regular part of developmental literature only in the 1980s. Despite this long history, or perhaps because of it, competing definitions of capacity abound, muddying the theoretical waters. Multiple meanings have developed because of the concept's use across varied disciplines, including politics, history, sociology, and economics, not to mention its use in applied development policy. The capacity concept is also used across a variety of cases and levels of analysis, further adding to variety of meanings. The lack of a coherent, consensual definition is evident in even a cursory review of the literature. What follows is a summary of major theoretical orientations in the capacity literature and example empirical studies following from those orientations. These orientations and historical and quantitative examples are summarized in the table below. (For a similar summary of capacity literature grounded in security and conflict studies, see Kocher, 2010).

Early capacity concepts were rooted in a fundamental conflict between the state and civil society. Capacity theories that argue for state autonomy or for state power view capacity through the lens of state against civil society or state against civil society and other states. Other theories have not viewed state capacity as a function of conflict, but instead see capacity as a function of policy. These theories see state capacity through the lens of policy preferences, decision-making, and implementation, with an emphasis on technocratic competence. While the earlier, state-society oriented definitions of capacity lend themselves well to national and cross-national analysis and comparisons, the policyoriented definitions lend themselves to national and sub-national analysis and comparisons. We provide a brief account of these various concepts along with empirical examples that utilize quantitative data or historical comparative illustrations.

Weberian notions of autonomy and bureaucracy informed early definitions of state capacity. Skocpol and Finegold (1982) have argued that state strength comes from autonomy from civil society and its power holders. Later, Skocpol developed the Weberian concept of the state further to argue that state capacity is a function of state autonomy, integrity, bureaucratic refinement, and resources (Skocpol, 1985). While other definitions of capacity have come to dominate thinking (see below), state autonomy nevertheless remains a part of theoretical and empirical work. Examples of empirical work following the concept of state strength as state autonomy include Skocpol and Finegold's (1982) historical account of New Deal-era federal economic interventions in industry and agriculture and Doner's (1992) study of Southeast Asian auto industry development.

Capacity as autonomy has been challenged primarily through variations on the concept of capacity as power. Migdal (1988) defined capacity as "the ability of state leaders to use the agencies of the state to get people in the society to do what they want them to do" (p. xiii). This was taken to an international level by Kugler and Domke (1986), who defined power in international politics as "the ability of one nation to exercise control over the behavior or fate of another" (p. 39). These definitions are essentially Dahl's (1957) concept of power used in state-society level terms rather than at the individual level of analysis. Thus, the developments and criticisms of the concept of power stemming from this tradition can be applied here (see, for example, Bachrach & Baratz, 1975; Lukes, 2004). In addition to Kugler and Domke's (1986) study of war-time capacities, a fascinating application of this approach can be found in Gaventa's (1980) historical study of Appalachian miners and coal companies and unions.

Another line of theorizing in capacity literature has focused on the scope or range of a state's power. Mann (1984) argued that "despotic power" is "the range of actions which the elite is empowered to undertake without routine, institutionalized negotiation with civil

society groups" (p. 188). Here, range is curtailed by civil society. An alternative take on scope is exemplified by Fukuyama (2004), who argues that scope "refers to the different functions and goals taken on by the government" (p. 7). That is, the scope of a state's power is, at least in part, a decision made by the state itself. Thus redistributive activities of the state, such as welfare programs, are variations in chosen scope rather than capacity itself. The notion of range, similar to autonomy and power, focuses on the negotiations within the state and between it and other actors regarding level, type, and form of intervention in society.

Exploiting the distinction between capacity and scope, Centeno and Portes (2006) focus their attention on the interaction between the two—what they call "regulatory intent". The results of state policies will be a product of both what state seek to accomplish (scope) and what they are able to implement (capacity). For them, the critical theoretical category is that of the "frustrated state": those wishing to impose much greater control over a society than they are organizationally capable of doing so.

Other definitions have taken a narrower approach, arguing that capacity is the ability of the state to form a policy decision and implement it. An example comes from Mann (1984), who defined infrastructural power as "the capacity of the state to actually penetrate civil society, and to implement logistically political decisions throughout the realm" (p. 189). This is a significant conceptual narrowing. Similarly, Fukuyama (2004), defines capacity as "the ability of states to plan and execute policies and to enforce laws cleanly and transparently" (p. 4). An example of empirical research in this line of capacity concept includes Geddes ([1994] 1996) study of politicians and policy in Latin America.

Note above that Fukuyama adds the normative qualifiers that state actions should be carried out "cleanly". This normative approach is one which is common in development practitioner literature. We argue that capacity is not and should not be a normative concept. That is, whether a state uses its capacity to enact policy preferences for "good" or "bad" ends does not negate the empirical reality of that state's ability. Historical examples of Soviet Russia and Nazi Germany during the World Wars are obvious examples of states with significant capacity (in any of the above conceptions) which was used in morally deplorable ways.

We add here that the study of political elite decision-making is an understudied area, even as it is pointed to as important in the literature. Geddes ([1994] 1996) argues, "If one wants to explain a state's preferences regarding development strategies, for example, one needs to know who has power and what they want and believe" p. (6). Despite this insight, elite decision-making in the developing world is still a neglected area of study.

Finally, capacity is often been understood in a dual relationship to wealth: both how much wealth a polity produces (GDP) and how much of that wealth can be extracted by the state (taxation). State capacity as a function of wealth (and wealth extraction) was formalized theoretically and historically for Western countries by Tilly (1985). The influence of this idea, even if limited to state intervention and development literature, is difficult to overstate. Theoretical development of the idea is demonstrated by Levi (1988). Recent empirical studies directly related to capacity and taxation include Besley and Perrson (2009). Empirical studies using taxation as a measurement of capacity are myriad (an example includes Shen & Williamson, 1997). Nevertheless, the association between state strength (and state building) with wealth and wealth extraction is not without problems.

Undoubtedly wealth of a country (GDP) and access to that wealth by the state (taxation) are major components of a state's ability to develop in ways that are important in the context of state building. That said, GDP and taxation have come to be seen as "best measures" of capacity. We argue that this is problematic for three reasons. First, focusing on wealth of a country limits the capacity concept in unnecessary ways. It leads to emphasis of classic Western industrial meanings of development, which may not be the best means of state building in non-Western developing countries. Second, despite the arguments presented by Levi (1988), variations in taxation levels are not explained well simply by a concept of "capacity". It is clear that among developing and developed countries, the variations in taxation are as much about a complex understanding of national history as an understanding of capacity, however defined. Finally, we want to further the idea that development and state building can occur at many points on the GDP/taxation curves. Wealth itself is not enough of an explanation; we want to understand the variations in development that are occurring at similar income levels.

Table 1.

Capacity Concept	Source	Empirical Examples
Autonomy	Skocpol (1985)	Skocpol and Finegold (1982), Doner (1992)
Power	Mann (1984), Migdal (1988)	Kugler and Domke (1986)
Scope	Mann (1984), Centeno and Portes (2006)	
Policy Decisions/ Implementation	Mann (1984), Fukuyama (2004)	Geddes (1994)
Wealth (GDP/ Taxation)	Tilly (1985), Levi (1988)	Shen and Williamson (1997), Besley and Perrson (2009)

Summary and Capacity Concepts, Sources, and Empirical Examples

Complicating the Capacity Concept

There are a number of continuing problems that are raised in this review of the concept of capacity. First is the potentially devastating threat of tautology that permeates so much work on capacity and strength. For example, Huber (1995) argues that a primary goal of states is enforcement of rule of law, while also arguing that rule of law is a necessary but not sufficient component of the achievement of other goals. Here she has already blurred the line between inputs, goals, and the ability to transform one into the other. The threat of tautology creeps in particularly during attempts to operationalize the concept for empirical work. It is simple enough to conceive of capacity as variable, with states having more or less capacity at any given time in any given area. However, upon measurement, often the concepts and variables that are used to define capacity (institutions, bureaucracy, infrastructure) are the very same concepts and variables used to determine the outcomes of capacity. How does one separate the measurement of capacity from the measurement of the *results* of capacity?

Related, there is inherent in observational measurements of capacity the problem of sampling on the dependent variable. That is, the very states that are at a minimum level of capacity survive to be measured and have the ability to measure themselves (for example, provide statistics on mortality and literacy). Indeed, perhaps an underutilized "measure" of capacity is simply a dichotomous measurement of whether the data exists for a particular variable. One relatively simple strategy in this line might be to measure the regularity with which statistical digests and the like are produced.

A third issue, and completely ignored in the literature, is the variation in results which simply reflect the variation in ideological priorities of different cultures and social groups. For example, rates of female literacy may reflect capacity to educate a population, but also the extent to which parts of a society oppose such efforts, the strength of their resistance, and dispersion of population which might limit the reach of universal education. More importantly, measures of results fail to differentiate between the capacity of the state to implement policies, the legitimacy of these policies, the capacity of a society to resist implementation, and the simple natural obstacles involved. We argue that only qualitative and historical work can fully address these sources of variation in capacity.

The Problem of Measurement

The problems of measuring capacity come from two areas in particular. The greatest source of measurement problems stem from the theoretical problems of capacity. Because the concept of capacity has yet to be fully developed, it is no surprise that operationalization and measurement of the concept have also yet to be convincingly developed. For example, the most common measure of capacity is tax revenue, though it is sometimes considered a measure of capacity (for example, Kugler & Arbetman, 1997; Kugler & Domke, 1986; Shen & Williamson, 1997;

Tilly, 1992) and other times a measure of scope (for example, Fukuyama, 2004).

Competing definitions and understandings of the concept have led to competing measures. For example, Geddes ([1994] 1996) argues that the ability to tax and coerce private actors depends on the existence of effective bureaucratic organizations (p. 14). Therefore, she conceptualizes capacity indirectly through bureaucratic autonomy and organization, operationalized as votes by party and seats in the legislature. Alternatively, Doner (1992) follows in the tradition of Skocpol (1985; & Finegold, 1982) and operationalizes capacity based on the state's autonomy from civil society. Capacity is often measured through composite scores and indices, assuming that a complex of factors combine in a particular way to indicate state capacity. From these assumptions come measures of human development (the UN's Human Development Index (HDI)), competitiveness (the World Economic Forum's Global Competitiveness Index), and governance (the World Bank's Governance Indicators). An interesting critique and response exchange on the World Bank governance measures was published in the May 2007 issue of The Journal of Politics (Kurtz & Schrank, 2007b; Kaufmann et al., 2007c; Kurtz & Schrank, 2007a; Kaufmann et al., 2007b). Clearly a lack of concept development and consensus has led to a wide variety of measures which often point in conflicting directions.

A second source of measurement problems is data quality, which has proved to be a hindering factor in analysis. Data quality is threatened by both availability and objectiveness. Even when an operationalization has been decided upon, such as tax revenue as a percent of GDP or police force as percent of the population, there is a disappointing lack of data available. In our review of 45 variables related to capacity for this paper, 18 out of 45 (40%) developing countries were missing 5 or more variables; 5 countries were missing more than 10 variables, including Saudi Arabia (missing 17), United Arab Emirates (15), and China (13). Particularly in areas of governance and labor, such as voter turnout, what information that is reported is of questionable objectivity. For example, reports of voting age population turnout for the most recent national election met or exceeded the voting age population count in Angola (121%), Rwanda (93.6%), and Vietnam (101%).

Capacity vs. Utilization

One way to begin to overcome these theoretical and methodological problems is to rigorously attend to the separation of inputs and outcomes and the transformations from one to the other. We suggest that the applied economics literature has already developed a rich theoretical and empirical literature from which we can learn. This literature surrounds the concept of capacity utilization.

Capacity utilization was introduced early in the 20th century as part of economic studies of industrial output in the United States. Theoretical issues and measurements were discussed by Klein (1960) and Klein and Preston (1967). Morrison (1992) has provided a solid overview of both the theory and its many measurement variations. Berndt and Hesse (1986) have demonstrated that the capacity utilization can be used as a comparative international measure.

Lessons to be learned from this literature that can be incorporated into future refinement of the capacity concept fall into both theoretical and measurement areas. Theoretically, the basic distinction between capacity and capacity utilization is critical. In the economic literature, capacity utilization is defined as "a ratio of the actual level of output to a sustainable maximum level of output, or capacity" (Corrado & Mattey, 1997, p. 152). Capacity utilization is often expressed as $u = Y/Y^*$, where u is capacity utilization, Y is a measure of maximal capacity, and Y^* is some measure of actual capacity output.

Capacity is, at its heart, often a latent variable (Kugler & Domke, 1986). It is understood in the econometric literature that full utilization of capacity is the goal, though not often the outcome. They are careful to distinguish in their measures what is *actual* capacity vs. what is *maximal* or *optimal* capacity. It is the ratio of the two that is utilization. We would do well to consider in our theories of capacity how we can distinguish best between optimal capacity, actual capacity, and to be clear that what we are currently measuring is often *utilization* rather than capacity itself.

Further, the distinction between inputs and outputs has been made very clear in the capacity utilization literature. For example Klein (1960) measured output as a function of labor input and capital stock (p. 274). Even if measured dynamically as flows of labor and capital stock, outputs are still distinct from these inputs. Strong separation of inputs and outputs should be a goal of capacity studies.

Another important contribution is the understanding of short-run vs. long-run outputs and their relationship to measurement (see, for example, Berndt & Hesse, 1986). Short-run capacity is distinct from long-run capacity, thus measurement of these two concepts should also be distinct. Cross-sectional data captures short-run capacity. Long-run capacity, the ability of a plant or a country to sustain a given output (outcome) can be capture only through panel (repeated-measures) data. This seemingly basic measurement fact has rarely, if ever, been employed in empirical studies of state capacity.

The applied economic literature has worked hard to overcome the inherent difficulties of measurement of both total capacity and actual capacity. Academics concerned with capacity would do well to visit this literature. The industrial economic community has developed longrunning surveys of firms and individual plants which use individual expert responses to determine what maximal output and actual output are. Data quality and proxy measures have certainly been at the fore of econometric literature, and it is foolish to ignore the advances made in the economic community to overcome these problems.

The Problem of Wealth

One of our central concerns with measures of state capacity is that they tend to reflect performance or delivery of service. In and of itself this should not be a problem, except when such indicators do not accurately reflect the institutional capacity of the state, but extraneous factors such as policy preferences or societal responses. We are particularly concerned with measures of state capacity that essentially mirror wealth. That richer states may be better able to deliver better services than poorer ones should come as no surprise. We contend that in the case of high correlations with wealth, these indicators might be of limited value as they do not provide any insight into the relative capacity of the state to work with the resources at hand. Obviously, these relationships might also reflect the positive effect of state capacity on development, but in the absence of careful longitudinal study, comparative statistics might also be susceptible to spuriousness.

Perhaps the most egregious case of this is the infrastructural index provided by the World Economic Forum which is significantly correlated with all measures of wealth from total GDP (0.78) to either measure of per capita income (0.74 and 0.53). This statistical relationship leads us to question the utility of depending too much on reputational surveys where respondents are asked to rate the quality of their governmental services with an appropriate comparative context. When the sample is drawn from a sector of the population that has had access to the quality of services (in this case, infrastructure) of the developed world, the responses are even more problematic. Why would we expect a middle-income state to have the same level of highways as the wealthy, no matter the quality of governance? Similarly, where we find relatively strong correlation between World Bank Indicators and GDP per capita, both as measured in US dollars (USD) (0.57 to 0.62) or in PPP (from 0.64 to 0.68) as well with HDI (0.53 to 0.62). Again, in such instances, what appear to be relatively effective states may simply be rich enough as to allow an element of inefficacy while still performing at acceptable levels.

Another possible candidate for measuring state capacity which has received some attention is the informal sector. Both of our measures of informality (as reported by the ILO and the OECD) are even more highly correlated to per capita income (-0.62 and -0.68 as measured by \$US and -0.74 and -0.63 by PPP). The relationships here are of course complicated by the fact that poorer economies are more likely to generate these types of jobs regardless of the capacity of the state to regulate them.

Interestingly, we find considerable variation in the extent to which some of the individual (and non-reputational) infrastructural measures are also highly correlated with wealth. On the higher end, delivery of infant and maternal health (as measured by mortality rates), water and sanitation infrastructure, and secondary education appear to be correlated with income (-0.4, -0.37, 0.40, 0.41, and 0.48 for USD and -0.52, -0.49, 0.5, 0.53, and 0.62 for PPP respectively). Note that such levels would still indicate considerable analytical space for the importance of institutional capacity in determining delivery of services.

The much lower level of correlation with primary enrollment (0.22 and 0.31) indicates that this has become so universally expected of even relatively poor societies that variance has declined precipitously (this is also true of measures such as immunization).

Two measures that appear surprisingly free of an income effect are quality of the road network and postal delivery. We suggest that much more work could be done on analyzing the institutional bases of these two aspects of good governance. Similarly, voter registration and turnout are uncorrelated with GDP per capita as well as tax revenue as a proportion of GDP. Violence (as measured by the homicide rate) appears to be also independent of income for this group. Interestingly, we find no immediate relationship between the size of the state (as measured by the proportion of the population working in the public sector or as the percentage of government expenditures within GDP). Nor is the state's extractive capacity, as measured by tax revenue as a percentage of GDP and percentage of the population in the armed forces.

Given the importance of comparative wealth in these outcome measures, we are particularly interested in identifying those countries that do better or worse than might be expected given their income levels. That is, controlling for income (but again, not for many of the other possible influencing factors) what states seem to deliver more or less? We may begin with the World Bank indicators. Not surprisingly given the high correlation between these measures, the same set of countries is found to over and underperform. The star in the former category is certainly Chile. South Africa, Malaysia, and the Baltic republics also perform above the expected income line. Among underperformers we find two different set of countries: African states and those dependent on primary resources.

In terms of extractive capacity, we do not find such clear patterns. Obviously, the depth and reach of conscription will be highly correlated with geopolitical context and most of the countries with high percentage of the population in the armed forces find themselves in permanent states of war preparation: the Middle East, Taiwan, and the ROK. For our measure of tax, complications include differing accounting methods and levels of extraction responsibility (e.g. national vs. provincial) and the avail-ability of other sources. Cases of interest here include South Africa, Algeria and Morocco for above the line and Bangladesh and Pakistan below it. If we analyze the delivery of a set of government services controlling for income we find two pat-terns. The first is a significant clustering of cases close to the mean with few if any standouts for any particular measure. The two exceptions may be secondary enrollment and provision of sanitation where we believe we may speak of institutional legacies. The best performers in these measures seem linked by the relatively long-term existence of a state apparatus that built the underlying foundations for service delivery and this includes the ex-Soviet republics and some of the wealthier Latin American states. The worst performers (and here the variance is much greater) seem to be consistently African countries.

There has been considerable discussion regarding the link between state capacity and integration with international economy. On the one hand, some argue that the need to produce for an export market would lead to a strengthening of the state as it sought to improve both its infrastructural base and its native human capital. On the other, discussions of the resource curse would lead us to expect relatively weak performance as there were few incentives for the state to develop institutionally given the (relatively) easily available bonanza from exploiting a primary resource. Surprisingly, our preliminary analysis does not support either position. No single measure of state capacity has a significant correlation with the traditional measures of the phenomena discussed above (exports as a percentage of GDP for the first and primary exports as a percentage of GDP for the second appear to contribute in any consistent fashion to the delivery of state services.

The Multiple Applications of Capacity

One of the most disturbing findings in our review of measures of capacity is how little any particular measure relates to any other. Our data set reviews 80 variables across 45 countries (Centeno, 2012) selected for their status as developing nations of some recognizable capacity. The country list was selected based on middle rankings of human development, specifically avoiding countries of high income and very high development and particularly low income and very low development. Pairwise correlations of all variables were run in order to

account for missing data in the set. There were few significant correlations among the variables. Most correlations are below 0.30.

Expected significant correlations include The World Bank Governance indicators (effectiveness, rule of law, and control of corruption) among each other and between the World Bank Governance Indicators and wealth. Total literacy and female literacy are highly correlated, as are HDI scores and literacy, infant and maternal mortality, and access to drinking water. These expected and significant correlations attest to the accuracy of the data set.

However, there are nearly no other statistically significant correlations other than those between wealth and human development and between wealth and infrastructure. This is in stark contrast to the findings of Holmberg et al. (2009). Holmberg and colleagues find significant correlations between the three World Bank Governance indicators and societal outcomes. However, it must be pointed out that they use a data set that includes up to 180 countries. Given that their data set includes the most highly developed countries (and wealthiest) as well as the least highly developed countries (and poorest), there is no doubt that they would find significant correlations. Our data set analyzes specifically moderately developing countries, seeking to understand why countries of comparable income or historical situation have varying measured outcomes.

As has been argued by Migdal et al. (1994), it is critical to understand that states have greater or lesser capacity across any number of areas. Indeed, much earlier Skocpol and Finegold (1982) argued that the United States federal government displayed a particularly strong capacity to form and implement a development program in agriculture during the New Deal era, but was incapable of developing and implementing a similarly successful one in the industrial sector during the same period. This is an excellent example of how a particular state can have varying capabilities across sectors even at the same historical era at the same level of aggregation. The lack of clustering among the outcome variables points to the importance of understanding the variability of capacity within countries and regions. The empirical evidence suggests that capacity outcomes are generally not inherently related to each other. Wealth obviously accounts for variation in development. However, accounting for that relationship, there is not compelling empirical evidence that strength in a particular sector translates necessarily or even easily to strength in another. This fact has yet to be explored, much less explained.

The Rule of Thirds: The Importance of Regionalization

An interesting problem with capacity is that as it stands, it does not contextualize outcomes in any way. Examining various measures of development by regions shows how important contextualization can be. Our empirical analysis finds important variation across regions and within regions, which are both totally obscured by capacity measures as they are currently used. Further, we find that running simple regression analysis on our outcome variables, accounting for regionalization explains typically a third or more of the variation.

Two examples illustrate the importance of understanding variation across regions. First, an analysis of access to sanitation as a function of income was run. World Bank data was used to plot access to improved sanitation facilities as a function of GDP/capita in year 2000 USD. A basic scatter plot of these two variables indicates positive linear relationship (see Figure 1). If we plot the data with regional indicators, however, it appears that regional grouping is high (see Figure 2).

Two separate regressions were run, first with sanitation on income then with sanitation on income with regional controls. The regression estimates indicate that income is a significant positive predictor of access to improved sanitation in both models. However, the R^2 (goodness of fit estimate) in the first model (without regional controls) is 0.279. That is, income as measured by GDP/capita (USD) alone explains 27.9% of the variation in the model. However, including controls for regions increases the goodness of fit further to 0.607.

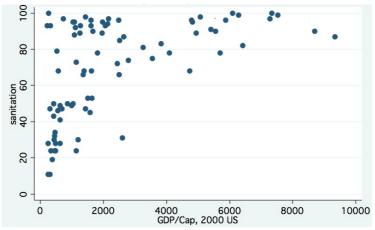


Figure 1. % Population with access to improved sanitation by GDP/capita, 2000 USD

Source: World Development Indicators, The World Bank

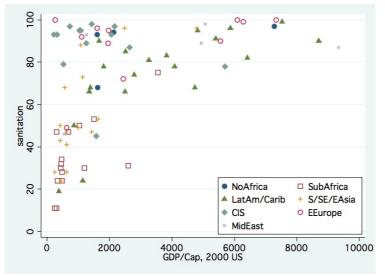


Figure 2. % Population with access to improved sanitation, regionally marked by GDP/capita, 2000 USD

Source: World Development Indicators, The World Bank

Incredibly, almost a half of the remaining variation (that is, an additional third of total remaining unexplained model variation) is explained by accounting for region alone (See Table 2 for estimates).

	(1)	(2) Sanitation
	Sanitation	
GDP/capita (USD, 200)	0.00645 ***	0.00466 ***
	(5.72)	(5.72)
North Africa		42.53 ***
		(4.35)
Latin America		26.84 ***
		(4.24)
S/SE/E Asia		19.19 **
		(3.10)
Post-Soviet		49.43 ***
		(7.66)
East Europe		43.59 ***
		(6.15)
Middle East		32.12 ***
		(3.44)
Constant	54.05 ***	30.75 ***
	(14.88)	(6.88)
Ν	83	83
adj. R ²	0.279	0.607

Linear regression results for sanitation

Table 2.

t statistics in parentheses

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

150 Enriquez & Centeno - State Capacity

Plotting the regression lines by region highlights the variation across our data set (See Figure 3).

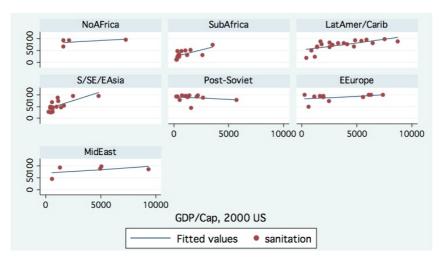


Figure 3. % Population with access to improved sanitation by region with fitted lines, by GDP/capita, 2000 USD

Source: World Development Indicators, The World Bank

Running similar regression analyses on the outcome variables of maternal mortality showed the importance of regionalization to an even greater degree. The goodness of fit (adjusted R^2) score for maternal mortality on GDP par capita (2000 USD) is 0.198. That is, 19.8% of the variation in outcomes is explained by GDP per capita. However, when we control for regional groupings, that goodness of fit score jumps to 0.767, or 76.7% (see Table 3).

Time and again we found that accounting for regional groupings improved our analyses by a third or more. It is clear that accounting for geographical context is critical to understanding development and capacity outcomes. We also want to note that this geographic importance points to the need for historical comparative models and explanations.

	(1) Maternal Mortality	(2) Maternal Mortality
GDP/capita (USD, 200)	-0.0433 ***	-0.0197 **
	(-4.66)	(-3.31)
North Africa		-436.3 ***
		(-6.93)
Latin America		-424.1 ***
		(-10.49)
S/SE/E Asia		-341.4 ***
		(-8.40)
Post-Soviet		-530.8 ***
		(-12.75)
East Europe		-514.6 ***
		(-11.50)
Middle East		-444.2 ***
		(-8.10)
Constant	291.3 ***	600.5 ***
	(9.20)	(20.87)
Ν	85	85
adj. R ²	0.198	0.767

Table 3.Regression results for maternal mortality

t statistics in parentheses

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

Within Region Variation: When Money Doesn't Matter

Variation within regions is equally important. Given the shared history, resources, and social and cultural factors of particular regions, basic capacity measures as they currently stand do not explain within region variation. When plotting homicide rates as a function of GDP/capita (2000 USD), it is clear that there are outliers in the data (see Figure 4).

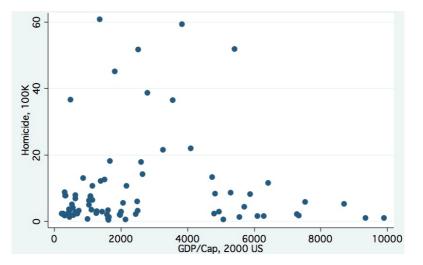


Figure 4. Homicide per 100K by GDP/capita, 2000 USD Source: UN Office on Drugs and Crime

Plotting the data with regional labels, we see that 1) the outliers are from primarily Latin American countries, and 2) what seems to be a strict relationship between income and homicides is actually highly regionalized (see Figure 5).

While the Latin American countries have highly varied GDP per capita numbers there is limited relationship between increase in income and decrease in violence (as operationalized as homicide per 100K). Honduras has a GDP per capita (2000 USD) at 1352.79 and very high homicide levels. Venezuela has considerably higher GDP per capita at 5401.02, yet only limited declines in homicide. In a reverse situation, Sub-Saharan Africa clusters strongly around a similarly low GDP per

capita, though there is wide variation among their homicides rates, particularly with South Africa with the highest homicide rates of all.

Capacity theory or measures currently do not explain these crossregional and intra-regional variations. While some regional clustering is to be expected due to contagion effects, the extreme results found in the most basic data analysis are not accounted for by current theory.

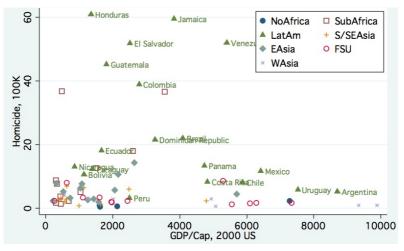


Figure 5. Homicide per 100K by Region by GDP/capita, 2000 USD Source: UN Office on Drugs and Crime

An Empirical Example: Durable Capacities and the Case of Literacy in the Former Soviet Union

We have argued that capacity as a concept has been defined variously in the literature. Major problems of all definitions have been a lack of separation between inputs and outcomes. Further, wealth has been used as a proxy for capacity to the detriment of other, more robust explanations, such as regional influence and (relatedly) historical explanations. Finally, empirical studies have relied heavily on crosssectional data, giving virtually no insight into long-run capacities of states. Here we present a simple analysis of literacy levels in former Soviet Union (FSU) countries. We use FSU countries from our data set to argue that a particular capacity outcome, adult (15+) literacy levels of a country over time, have remained highly stable despite considerable variation in per capita income both within and across countries. This stability points to a high level of capacity by the state to provide (at least) a basic level of education to its people despite wealth variations. We argue that this can only be explained through historical analysis.

Our analysis looks at 13 former Soviet countries: Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. The data consist of adult (15+) literacy as reported by the World Development Bank between 1985 and 2005. All years that had data were used. This resulted in at least two data points for all countries, several with at least four.

The former Soviet Union countries were chosen to highlight our general arguments presented thus far. First, it is possible to separate investments into development outcomes from the outcomes themselves. Second, wealth does not explain all variation of a particular outcome—the context of region and history must be employed to fully understand outcomes. Third, these cases highlight the durability of capacity investments over time and across wealth variations.

It is well known that the Soviet Union made literacy and broader cultural education a priority during its existence, and this priority was realized in a number of investments (inputs) to the goal. The investments included construction of schools throughout the vast realm of the Union, development and publication of standardized learning materials, and training and deployment of educators at pre-primary, primary, and vocational levels all over the country. These capital and labor investments were bolstered by the infamous propaganda machine, inculcating the importance, even the necessity, of literacy and cultural awareness to the ideal Soviet person. The outcome of this investment was a skyrocketing literacy level from pre-Soviet to Soviet eras. At the time of the dissolution of the Soviet Union, our 13 countries showed adult literacy levels of 95% or greater.

What cannot be captured by simple cross-sectional data is that extreme variation in per capita income that the FSU countries have experienced, both within their own boarders over the past 25 years and across the various countries. The graph below (Figure 6) demonstrates the variation income experienced by these countries. Reviewing the graphs, two things should become clear. First, there is variation across countries in per capita income. Kazakhstan has had considerably higher income levels across time than its neighbors Kyrgyzstan, Tajikistan, and Turkmenistan. Similarly, Russia has had higher income levels that its neighbor Ukraine. Second, across time, each country has had considerable variation in the 25 years shown in the graphs. The history of the dissolution of the Soviet Union was also a history of financial crisis, hyperinflation, and then increasing and steady economic improvements. This is demonstrated most clearly by countries such as Belarus, Georgia, and Latvia.

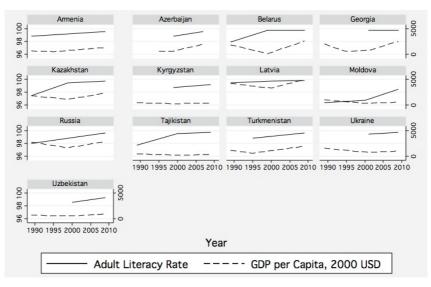


Figure 6. Adult literacy and GDP/capita, 2000 USD Source: World Development Indicators, The World Bank

Despite these variations in wealth, both within and across countries, literacy does not decline, and in fact only improves over time. In every case shown, literacy begins at a high level and maintains or increases levels. This demonstrates the *durable capacity* of education and literacy in these nations. Considerable investment was made during the Soviet era to create education as a cultural norm and to build up infrastructure to educate the ever-increasing millions of Soviets. This capital investment proved to be durable across the various levels of income of the member countries during the Soviet era. It also proved to be durable to wealth and social changes after the dissolution of the Soviet Union.

Capacity as a Relational / Contextualized Concept

Although a small number of theorists have called for an understanding of capacity that includes contextualization (Huber, 1995; Migdal, 2001), nothing has been done to fully develop such a theory, and nothing has been done to demonstrate what such a theory would lead to in terms of empirical research. There are a number of areas that require further development immediately. These include national as well as subnational comparisons, geographic and regional variations, and historic contextualization. Most needed is serious attention to historical case studies based on empirical data from a number of years, regions, and levels of aggregation.

Both the theoretical and empirical issues of state capacity point out the difficulties of exploring what remain institutional "black boxes." Social science has been very successful in measuring and analyzing the results of state policies such as stages of development, degrees of democracy, and levels of human welfare. We know much less about what goes on inside the states. This is particularly true of sub-elite practices. For example, we certainly know much more about the technocrats of central banks and finance ministries than about revenue collectors in the provinces. Studies of the professionalism and skills of civil service tend to be country specific and, apart from the groundbreaking study by Evans and Rauch (1999), we have little by way of concrete measures or indices across enough cases for meaningful comparative analysis.

We need to specify the geographical diffusion of both state intentions and implementations. With regards the first, some states may be only concerned with policies in particular locales (e.g. major cities) or sectors (taxation or conscription vs. economic development). Similarly, the authority of a state maybe geographically bounded. In the most extreme cases, state authority may be limited to a few blocks around a presidential palace. In some circumstances. intention and implementation combine to limit the reach of a state. So, for example, the apartheid South African state was much more concerned and capable of dealing with criminal activity in white areas than in those defined as "African." In general, we might say that uniformity of reach may be an excellent measure of general state capacity.

Finally, no analyses we have seen account for historical variations in capacity, either within a particular country over time or across a number of countries over time. This lack of inclusion of history in basic assessments cannot be overlooked. Simple single figures, such as annual rate of growth can be included in analyses to exploit changes over time. For example, the World Bank's Governance Effectiveness score correlates highly with GDP/capita (PPP) in our data set at 0.684 (at greater than 0.01 significance), yet has no significant correlation with annual growth for the same group. Thus, as a measure of capacity to change, income per capita has no value.

We have presented a number of concepts for development in this paper, establishing an on-going research agenda. We have demonstrated the long-standing use of the concept of state capacity and also its many definitions and operationalizations. It is clear that further research in the area requires the establishment of clear theoretical distinctions between the independent and dependent variables related to state capacity and clear distinctions in their measurement.

We have also established the idea of *capacity utilization* and its relevance to state capacity research. Industrial economists have developed both theoretical and empirical research related to capacity utilization, and it would only benefit the fields of comparative politics and development scholars to apply those insights to state capacity. A key contribution of capacity utilization will be to understand which states are operating a high capacity and which have yet more resources, in all their forms, to devote to services and development for its citizens.

Capacity utilization will also benefit development scholars by pointing us to states (and sub-state entities) that are "doing more with less." That is, actors that are providing relatively high levels of services and goods with what limited resources they have, primarily through efficiency.

Next, and very much related, we have established the limited value of GDP/capita as an explanatory variable for state capacity. There is no argument that variations in wealth are related to development outcomes. However, as our initial analyses have shown, the amount of variation explained by wealth in this group of countries -those that are neither at the very top of the UN Human Development Index nor at its very bottom- is limited. Wealth of a country per capita can explain 30% -40% of variation. While we do not want to ignore this, it is clear that wealth does not explain all of development. We feel strongly that economic wealth and growth has its place in the literature, but more fruitful research will focus on the other explanations for variation. We suggest here that using regional variation, a proxy for historical context, is one important step forward. We also suggest that sub-national research is required to further understand the role of the state in development. Critically, the answer to the question, state capacity for whom? will be answered through this kind of research.

We have also introduced the idea of *durable capacity*. This concept is critical to understanding the best policies in which a community can invest. Our basic example of adult literacy in the former Soviet countries demonstrates that lasting gains can be made in development outcomes *regardless* of per capita income. Further, these gains are not lost over time despite regime change and income variations.

As our main argument at the beginning of the paper is that the concept of state capacity is muddled, we are hesitant to add to that muddle. Nevertheless, we believe that the contributions of this paper will lead to an opening of the "black box" of the role of the state in development outcomes and more fruitful research into alternatives to wealth as the answer to capacity development.

References

- Bachrach, P., & Baratz, M.S. (1975). Power and Its Two Faces Revisited: A Reply to Geoffrey Debnam. *The American Political Science Review*, 69(3), 900–904.
- Berndt, E.R., & Hesse, D.M. (1986). Measuring and Assessing Capacity Utilization in the Manufacturing Sectors of Nine OECD Countries. *European Economic Review*, 30, 961–989. doi: 10.1016/S0014-2921(86)80001-0
- Besley, T., & Perrson, T. (2009). The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review*, 99(4), 1218–1244. doi: 10.1257/aer.99.4.1218
- Centeno, M.A. (2002). *Blood and Debt: War and the Nation-State in Latin America*. University Park, PA: Pennsylvania State University Press.
- Centeno, M.A. (2012). *Appendix to Enriquez and Centeno, RIMCIS, 1, 2, 2012.* Retrieved from http://scholar.princeton.edu/cenmiga/publications/appendix-enriquez-and-centeno-rimcis-1-2-2012
- Centeno, M.A., & Portes, A. (2006). The Informal Economy in the Shadow of the State. In P. Fernandez-Kelly & J. Shefner (Eds.), Out of the Shadows: Political Action and the Informal Economy in Latin America (pp. 23–48). University Park, PA: Pennsylvania State University Press.
- Collier, P. (2007). *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford: Oxford University Press.
- Corrado, C., & Mattey, J. (1997). Capacity Utilization. *The Journal of Economic Perspectives*, 11(1), 151–167.
- Dahl, R.A. (1957). The Concept of Power. *Behavioral Science*, *2*, 201–215. doi: 10.1002/bs.3830020303
- Doner, R.F. (1992). Limits of State Strength: Toward an Institutionalist View of Economic Development. *World Politics*, 44(3), 398–431. doi: 10.2307/2010544
- Evans, P.B. (1995). *Embedded Autonomy: States and Industrial Transformation*. Princeton, NJ: Princeton University Press.

- Evans, P., & Rauch, J.E. (1999). Bureaucracy and Growth: A Cross-National Analysis of the Effects of `Weberian' State Structures on Economic Growth. *American Sociological Review*, 64(5), 748–765.
- Fukuyama, F. (2004). *State-Building: Governance and World Order in the 21st Century*. Ithaca, NY: Cornell University Press.
- Gaventa, J. (1980). *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley*. Urbana, IL: University of Illinois Press.
- Geddes, B. ([1994] 1996). *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley, CA: University of California Press.
- Holmberg, S., Rothstein, B., & Nasiritousi, N. (2009). Quality of Government: What You Get. Annual Review of Political Science 12, 135–161. doi: 10.1146/annurev-polisci-100608-104510
- Huber, E. (1995). Assessments of State Strength. In P.H. Smith (Ed.) Latin America in Comparative Perspective: New Approaches to Methods and Analysis (pp. 163–194). Boulder, CO: Westview Press.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007a). Growth and Governance: A Rejoinder. *The Journal of Politics*, 69, 570–572. doi: 10.1111/j.1468-2508.2007.00552.x
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007b). Growth and Governance: A Reply. *The Journal of Politics, 69*, 555–562. doi: 10.1111/j.1468-2508.2007.00550.x
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007c). The Worldwide Governance Indicators Project: Answering the Critics. Working Paper 4149, World Bank Policy Research Department.
- Klein, L. R. (1960). Some Theoretical Issues in the Measurement of Capacity. *Econometrics*, 28(2), 272–386.
- Klein, L. R., & Preston, R.S. (1967). Some New Results in the Measurement of Capacity Utilization. *The American Economic Review*, 57(1), 34–58.
- Kocher, M.A. (2010). State Capacity as a Conceptual Variable. Yale Journal of International Affairs, Spring/Summer,137–145.
- Kohli, A. (2004). State-Directed Development: Political Power and Industrialization in the Global Periphery. Cambridge: Cambridge

University Press.

- Kugler, J., & Arbetman, M. (1997). Relative Political Capacity: Political Extraction and Political Reach. In M. Arbetman & J. Kugler (Eds.) *Political Capacity and Economic Behavior* (pp. 11-46). Boulder, CO: Westview Press.
- Kugler, J., & Domke, W. (1986). Comparing the Strength of Nations. Comparative Political Studies, 19, 39–69. doi: 10.1177/0010414086019001002
- Kurtz, M.J., & Schrank, A. (2007a). Growth and Governance: A Defense. *The Journal of Politics*, *69*, 563–569. doi: 10.1111/j.1468-2508.2007.00551.x
- Kurtz, M.J., & Schrank, A. (2007b). Growth and Governance: Models, Measures, and Mechanisms. *The Journal of Politics*, 69, 538–554. doi: 10.1111/j.1468-2508.2007.00549.x
- Levi, M. (1988). *Of Rule and Revenue*. Berkeley, CA: University of California Press
- Lukes, S. (2004). *Power: A Radical View* (2nd expanded edition). Hampshire, NY: Palgrave Macmillan.
- Mann, M. (1984). The Autonomous Power of the State: Its Origins, Mechanisms and Results. *Archives Européennes de Sociologie* 25, 185–213. doi: 10.1017/S0003975600004239
- Migdal, J.S. (1988). Strong Societies and Weak States: State–Society Relations and State Capabilities in the Third World. Princeton, NJ: Princeton University Press.
- Migdal, J.S. (2001). *State in Society: Studying How States and Societies Transform and Constitute One Another*. New York: Cambridge University Press.
- Migdal, J.S., Kohli, A., & Shue, V. (1994). Introduction: Developing a State-in-Society Perspective. In J.S. Migdal, A. Kohli & V. Shue (Eds.) *State Power and Social Forces: Domination and Transformation in the Third World* (pp. 1–6). New York: Cambridge University Press.
- Morrison, C.J. (1992). A Microeconomic Approach to the Measurement of Economic Performance: Productivity, Growth, Capacity Utilization, and Related Performance Indicators. New York: Springer-Verlag.
- OECD. (2006). The Challenge of Capacity Development: Working

towards Good Practice. DAC reference document, OECD, Paris.

- Przeworski, A. (1990). *The State and the Economy under Capitalism*. Chur: Harwood Academic Publishers.
- Shen, C.E., & Williamson, J.B. (1997). Child Mortality, Women's Status, Economic Dependency, and State Strength: A Cross-National Study of Less Developed Countries. *Social Forces*, 76, 667–694. doi: 10.1093/sf/76.2.667
- Skocpol, T. (1985). Bringing the State Back In: Strategies if Analysis in Current Research. In P.B. Evans, D. Rueschemeyer & T. Skocpol (Eds.) *Bringing the State Back* (pp. 3–43). New York: Cambridge University Press.
- Skocpol, T., & Finegold, K. (1982). State Capacity and Economic Intervention in the Early New Deal. *Political Science Quarterly* 97, 255–278.
- Soifer, H., & vom Hau, M. (2008). Unpacking the Strength of the State: The Utility of State Infrastructural Power. *Studies in Comparative International Development (SCID)* 43, 219–230. doi: 10.1007/s12116-008-9030-z
- Stiglitz, J.E. (2003). *Globalization and Its Discontents*. New York: W. W. Norton.
- Tilly, Ch. (1992). Coercion, Capital, and European States, A D 990–1992. Cambridge, MA: Blackwell.
- United Nations. (2000). *United Nations Millennium Declaration* number A/Res/55/2. United Nations General Assembly.

Elaine Enriquez is PhD Candidate in the Department of Sociology at Princeton University, United States.

Miguel Angel Centeno is Professor in the Department of Sociology at Princeton University, United States.

Contact address: Direct correspondence to the author at the Department of Sociology, Princeton University, Wallace Hall 116, NJ 08544, United States or at cenmiga@princeton.edu