Capturing State Strength: Experimental and Econometric Approaches

Capturando la fortaleza del Estado: Aproximaciones econométricas y experimentales

MARCUS J. KURTZ
Ohio State University

ANDREW SCHRANK
University of New Mexico

ABSTRACT

While almost all definitions of the state and state capacity (or stateness) derive at their core from a Weberian emphasis on the characteristics of institutions, the dominant measures of this concept rely instead on subjective, perceptions-based indicators. We propose an alternative strategy rooted in direct experimentation and econometric modeling to develop objective measures of the strength of state institutions, both for a contemporary cross section, but also (for the latter) for cross-time historical data essential for tracking the development of institutions.

Key words: State, State Building, Measurement, Objective, Experimental.

RESUMEN

Aunque casi todas las definiciones de Estado y capacidad estatal (o estatalidad) derivan del énfasis weberiano en las características de las instituciones, las mediciones de este concepto que predominan descansan, sin embargo, en indicadores subjetivos basados en percepciones. Este trabajo propone una estrategia alternativa basada en experimentación directa y modelización econométrica para desarrollar medidas objetivas de la fortaleza de las instituciones estatales, no sólo para un análisis cross-section contemporáneo, sino también (para la segunda) para datos históricos cross-time esenciales para rastrear el desarrollo de las instituciones.

Palabras clave: *Estado, construcción del Estado, medición, objetiva, experimental.*

* An earlier version of this paper was presented at the 'Stateness in Latin America' conference in Santiago, Chile March 2012. We'd like to thank (without implicating) participants at the conference and an anonymous reviewer for their comments and Kevin Arceneaux for a helpful email exchange. This article is part of the Millenium Nucleus for the Study of Stateness and Democracy in Latin America, Project NS100014, of the Ministry of Economy and Tourism of Chile.

THE CONCEPTUAL PROBLEM AND THE SHAPE OF A SOLUTION

It is customary in conceptual discussions of 'stateness' or 'state strength' to begin by defining the state through the invocation of Weber (1978: 54): a state is the territorial entity with the monopoly on the legitimate use of force. But this brief and intuitive definition can be misleading if it is understood in the framework of a classical approach to conceptualization. For Weber –though not necessarily for all those who have employed a Weberian approach to the definition of the state– these characteristics are clear and evident in the *ideal-typical* state; they are not strict "necessary conditions" for a territorial political entity to be a state. This is critical, for if we emphasize 'monopoly on the use of force' as a defining characteristic of statehood, this begs a major question: how is systematic inter-group violence within a territory overcome? If we follow North, Wallis, and Weingast (2009: 30ff.), the state emerges –and is in part defined by this role– as a mechanism that enables the control of otherwise severe inter-group, and especially inter-elite, conflict. Thus to define the state as a force monopolist (or at least a 'legitimate force monopolist') is to some extent conceptually truncate the sample: for it does not include those territorially-based institutional systems that have not yet fully established such a level of control (or societal support).[1] The second issue is the question of the "legitimacy" of a monopoly of force. This notion appears independently in the literature. For example, for Holsti, a coercively-powerful territorial entity is thus a state when its leadership is selected in accord with locally-relevant "ideas, doctrines, social ideologies, and/or constitutions that grant a comprehensive 'right to rule' to certain individuals or bodies" (Holsti, 2004: 30). And presumably it has greater 'stateness' to the extent that the belief in this legitimacy is more widely- and/or deeply-held in society. The third central feature of the state –one that is not directly referenced in this definition, but was of course extensively discussed by Weber– has to do with the organization of *administration*. States are institutions that use coercive and other resources to set rules and govern societies. And they vary quite substantially in how effectively they are able to do this.

The jumping off point for this paper begins with the potentially-provocative claim that neither of the first two elements of a Weberian ideal-typical notion of the state –the monopolization of force, or its legitimacy– should be used as defining conditions from the perspective of the assessment of 'stateness.' If, for example, we presume the monopolization of force, we are hard put to assess the institutions of the state where domestic challengers regularly impose threats to sovereignty.[2] But this forces us to omit from consideration the 'stateness' of, for example, Uruguay before the effective settlement of the armed struggle between Blancos and Colorados in the early 20th century, or the government of South Vietnam through 1975. But while neither 'state' came close to a monopoly on force, both were states, albeit comparatively weak ones precisely

1 Weber meant the ideal type to imply a monopoly on legitimate force - not necessarily all force. But in contemporary usage, both approaches appear.

2 Mann (1993: 55) similarly loosens this aspect of the Weberian definition, instead suggesting only that the state be "backed up by some organized physical force".

because they were unable to close off recourse to force as a means for civil society actors to challenge government. Legitimacy is an even more perplexing condition for the definition of a state, or an element in the conceptual core of 'stateness.'[3] To begin, as Przeworski (1986: 51) has pointed out, legitimacy is an inherently relational concept –and for illegitimacy to have a meaningful effect implies the existence of a preferable alternative. Thus two otherwise-identical political/administrative systems –including general societal approbation or disapprobation of their performance– could carry quite different levels of legitimacy merely because of variations in locally imaginable or viable alternatives. Moving this matter from difficult to intractable is the near-impossibility of measuring the legitimacy of a set of state institutions in a way uncontaminated by the level of satisfaction that respondents have with the outputs of public administration or the presence of alternatives.[4] However likely –and potentially important– state legitimacy might be in characterizing the highest levels of 'stateness', the concept has appears to have near-intractable problems of measurement.

The analysis here focuses squarely on the third dimension: the characteristics of the public administration. For here, we argue, can be found the range of variation that is crucial to understanding what Mann (1993: 59) calls 'infrastructural power' or what Huber (1995: 166) calls 'state strength': "the capacity to achieve the goals set by incumbents in chief executive positions".[5] There is still some slipperiness in this approach to conceptualization, because, as Soifer has noted (2008: 235ff.), as it has been used to understand state strength in terms of the *capacity* to undertake efforts as well as the *effects* the state has on society. While we agree with Soifer that an emphasis on capacity is crucial, even there we worry about conceptual conflation. For if capacity is about achieving goals or implementing decisions, this really has two components: the *material* resources to undertake the desired goal, as well as the *administrative* resources to do so.[6] For our purposes, the emphasis will be squarely on the latter: we want to be very careful to capture the principal/agent problems involved in the administration of the national territory, and not the comparative wealth of the society in which the administration is located.[7]

With this clarification in mind, we further refine this understanding of state capacity along what we think are the four most relevant dimensions. We begin with what we call

3 There is a case to be made that in principle the legitimacy of authority might be a conceptual component of stateness. But in a practical sense it poses severe measurement problems, and at the same time might equally be thought of more as a consequence of stateness than constitutive of it. For these reasons, we do not take it to be a defining characteristic.

4 Notably, Weber was well aware of this issue. As he notes (Weber, 1978: 214), "…the legitimacy of a system of domination may be treated sociologically only as the probability that a relevant degree of appropriate attitudes exist. *It is by no means true that every case of submissiveness to persons in positions of power is primarily (or even at all) oriented to this belief*".

5 Mann's more technical rendering focuses on the "institutional capacity of the central state… to penetrate its territories and logistically implement decisions".

6 We would raise also the possibility that a state might employ ideological resources that might similar affect outcomes. This possibility –however uncommon in a large-scale manifestation– necessarily poses challenges for output-based measures of stateness.

7 This is perhaps not so simple, since resources are themselves, to an extent, the product of administrative resources, most notably those of the taxation bureaucracy. This is all the more true if long-run economic development is itself a consequence of institutional quality as many, but not all, analysts, have suggested.

the *range* of the state. As Soifer (2006: 5-6) has importantly noted the ability of the state's institutions to reach the far corners of the national territory is essential, and in general state penetration of society is territorially uneven. Thus one aspect of the capacity of the state to administer society is the degree to which it can implement decisions across more versus less of the national territory. Our second dimension has to do with the *scope* of state authority; that is, over what institutional arenas is the state capable of making its decisions have effect? A state might, for example, be able to effectively police the hinterland and even impose some measure of taxation. But this does not imply that it is equally capable of providing basic public health, education for the population, or delivering the mail. Scope thus captures the idea of "strong at what?" in relationship to state capacity. Separate from range and scope, however, is what we call the *coherence* of the state. This has to do specifically with the principal-agent problems of decision-implementation. Here what is at issue is the degree to which line bureaucrats actually implement national-level policies effectively, and as written. That is, are bureaucrats performing their functions even where their actions may not be seen as beneficial. Are resources being distributed in a rule-bound fashion, or is corruption affecting the distribution of bureaucratic attention? This dimension thus includes both questions of human capital and ability of the bureaucracy itself, but also questions of the incentives facing bureaucrats in terms of meritocratic promotion, insulation from pressures from political actors, and a reward system that more generally directly relates to able accomplishment of tasks which bureaucratic agents are assigned. In sum, *coherence* should manifest itself in the uniformity of adherence to the terms set forth for decision making by the national executive.[8] Finally, we think it crucial to investigate the *temporal* dimension of state capacity. This reflects the degree to which a state's institutional capabilities are in fact institutionalized –rather than being dependent on, for example, the particular individual or partisan occupant of executive office. If, for example, effective implementation occurs because the bureaucracy is staffed by committed political cadre associated with a party or individual, it is unlikely to be durable, and might well manifest as inefficacy should executive office come to be controlled by a non-copartisan at a later point in time.

FROM CONCEPTUALIZATION TO MEASUREMENT

This conceptual rendering of stateness or state capacity or infrastructural power has from our perspective three principal merits. First, it is objective. That is, it does not rely on an effort to measure the beliefs of citizens about the nature of the state, the legitimacy of its leaders or the institutional procedures that selected them, or even perceptions of the efficiency of public bureaucracies.[9] Indeed, to the extent that these perceptions of or attitudes about the state are validly measurable, they are potential *outcomes* of state

[8] Related, but not as critical, is the question of the efficiency of bureaucratic action.

[9] There is obviously an immense literature that has developed and relied upon just such subjective indicators of governance, and our objections to it on conceptual and measurement grounds can be found in Kurtz and Schrank (2007).

capacity, not necessarily coterminous with it.[10] Second, and critically, they separate the question of policy from that of the institutions of public administration. This is what separates our approach to measuring the capacity of the state from those that emphasize, for example, objective measures of the extent of "red tape", the characterization of informants (firms, investors, consultancies) of the efficiency of bureaucracies, or even the level of security granted to property rights. All of these approaches intermingle policy evaluations with those of institutional effectiveness; for one person's red tape is another person's labor regulation, safety requirement, or environmental standard. Informants' evaluations of bureaucratic efficiency are shaped by a worldview that privileges the avoidance of regulation over its effective implementation. And even the level of security of property rights –and the distribution of whose rights are protected and how much– are importantly policy questions. Any number of very reasonable –and historically effective– economic-development strategies, environmental-protection systems, and labor-rights regimes have relied on the construction of far-from-absolute property rights.[11] Finally, our approach follows the recent trend toward the disaggregation of the notion of stateness –it assumes that states can be strong or weak in different parts of the national territory, with respect to different functional bureaucracies, or even at different points in time. It is our assumption that performance on different dimensions of state capacity may be correlated, but probably are not strongly so. And thus we must develop an approach that allows for the empirical evaluation of this possibility.

The problem of inference. The inference problem that most directly confronts us, however, is that in evaluating the capacity of a state administration to undertake policy, we do not always observe the underlying level of capacity. If, for instance, the pursuit of policy X requires a minimum level of state capacity Y to be undertaken efficiently, then what would be observable is Y. But if, for example, the state institution in question were capable of much more –but were not called upon to do so– we would wrongly conclude that its capacity is Y and not $Y + \varepsilon$, the true value. Capacity is a latent concept, and we can only directly observe an institution's capacity when it is called upon to do more than it is capable of or where differing levels of capability lead to differing levels of output, net of the policy goal. We turn next to efforts to solve this inferential problem in two very distinct settings - that for practical reasons call for the utilization of different approaches to measurement.

State Capacity in Small-N Comparative Analysis. In contexts where the goal is the comparative evaluation of a small set of states –or of a set of particular institutions with a state– we are most able to attempt direct measurement of the dimensions of state capacity. So, for example, we can begin to analyze the range of a state's institutions by examining the physical geography of its bureaucracy. What is the density in different subnational regions

[10] Indeed, the tendency for objective and subjective measures of state capacity to be at best weakly correlated suggests that conceptual conflation of these two issues is ill-advised.

[11] Indeed, instead of the literature's general characterization of property rights as 'strong' or 'weak', it makes much more conceptual sense to ask whether given a specific definition of the rights and obligations of property, how consistently and uniformly these rights and obligations are observed. But property rights are structured in quite distinct ways around the globe.

of line police officers, primary and secondary school teachers, or licensing or regulatory bureaucrats normed on population, school-age children, or commercial and industrial employment?[12] By comparing within and across functions in the subnational units, this approach would give us purchase on the reach and scope of state bureaucracies. The most challenging aspect of state capacity, however, is probably the most interesting: its coherence or effectiveness. It is one thing to know that there are, for example, a reasonable number of regulatory bureaucrats in a region. But this does not imply that they will effectively –and without graft– implement the policies set by their hierarchical superiors. Here we need direct evidence of a sort that the line bureaucrats have an incentive not to provide accurately (e.g., corruption, inefficiency, deliberate foot-dragging). We suggest two approaches, both experimental, for dealing with this problem:

1) Measurements of bureaucratic human capital and the incentive structures facing bureaucrats

In the literature on state capacity, efforts have been made to directly evaluate the incentive systems facing ordinary bureaucratic agents – often in the form of single country diagnostic exercises, or more promisingly, in a self-consciously cross-national comparative fashion (Rauch and Evans, 2000). This approach, however, relies on direct answers to sensitive questions from a small number of expert informants. We seek to improve sharply on this approach, by seeking similar information on meritocratic recruitment, comparative pay structures, informal payments, and illicit activity but focusing our efforts on a much larger, representative group of bureaucrats using tools designed specifically to elicit accurate information on inherently sensitive topics.

The generalized approach would be based on the survey list experiment, a technique pioneered in Political Science by Kuklinski *et al.* (1997) as a form of descriptive inference, later generalized to a multivariate framework by Corstange (2008) and Blair and Imai (2010). The merit of this tool is that it is by construction completely anonymous at the individual level, but nevertheless can reveal accurate aggregate information.[13] This technique could address questions like whether entry into the bureaucracy was really based on a competitive, meritocratic process; whether tenure protections are effective;

[12] To be clear of our inferential problem, however, we would want to make sure that the capacities observed are not associated with outcomes for which further improvement is theoretically impossible. These examples were selected precisely because it is unlikely that in most (at least developing-world) contexts that crime is largely controlled, literacy and basic educational achievement are effectively universal, and production is universally undertaken without substantial violation of prevailing tax, labor or environmental norms.

[13] The basic approach is quite simple: each respondent, randomly assigned either to a treatment or control group, is confronted with a list of three or four statements that are mostly comparatively innocuous (one sensitive statement is presented to the treatment group). The treatment group has a list that includes one additional statement, reflecting the issue of interest. For example, one might say "I am aware of employees in my Ministry who take payments in exchange for special treatment". The control statements are structured so that it is unlikely that anyone would agree with all or none of them. Then the respondents are simply asked with how many of these statements do they agree? The difference in means between the control and treatment group will then reflect the proportion of respondents who affirmed the sensitive statement, but without revealing any information as to responses on the individual level.

whether extralegal payments or extortion take place, and whether formal pay comprises most of a civil servant's income. This could be accompanied by direct survey information on less sensitive matters like age, educational attainment, pay grade, and years in service.

2) Measurements of the direct consequences of bureaucratic action

Measures of the incentive structure within which civil servants operate are only one of two general strategies that can be used to test public sector capacity. For incentives only translate into actual outputs in an efficient fashion if other –likely unknown or unmeasurable– characteristics of an institution do not impede its function. A second approach then would focus on the outputs of a public sector institution. The problem, of course, with output based measures is that they necessarily include information on policy choice and effort not just institutional efficacy. For if this is the case we run into the problem to which we alluded earlier. We would also add that this is a measurement strategy –and as such subject to error– and not a statement of the conceptual content of stateness.

To utilize this output-linked approach in a valid, comparative fashion, one should only examine public sector outputs that are not particularly politicized, and generally perceived to be essential state functions across a very broad set of states. Fortunately, there are quite a few examples of such activities. Most every state develops a system to deliver mail, most every major line Ministry has the function of providing important information (on law, policy, and procedure) to citizens, and countless regulatory agencies are charged with licensing and certifying a range of everyday activities (building codes, fire codes, occupancy permits, business licenses, etc.). Our approach is to utilize these omnipresent, everyday bureaucratic outputs to test the actual efficiency of public institutions. We explicitly avoid an emphasis on outputs that are at the center of political or policy debates, such as property rights, social policy, environmental regulation, or the like. It is a straightforward matter to design a system to test the efficacy and speed of mail delivery (intra-capital and inter-regional).[14] Similarly, for institutions of interest, application can be made (through all permitted means, written, electronic, by telephone, etc.) for information (on regulations, on how to comply, on where to submit documentation, etc.). The speed and correctness of the response would be very informative. Similarly, one could –pending some institutional review board considerations– even apply for the relevant permissions to start a minor firm and observe the consequence. How available is information on the relevant procedures? What are the processing times? Were any extralegal payments solicited? Basic tax collection efficacy –at least where a VAT is applied– could be ascertained by a sampling of purchases of standard goods in large, medium, and small retail outlets. The frequency with which, for example, proper receipts and cash register utilization are employed would be an indirect measure of the

[14] Indeed, Portes and Smith (2008) have used just such evaluations –based on expert informants, not on experimental investigation– of postal services and civil aviation authorities as metrics of state institutional capacity.

efficacy of a tax bureaucracy. Finally, striking at a core function of any market economy, one could also seek information on different types of real property ownership (land, firms, or homes) from the relevant functional equivalent of a Recorder of Deeds. These micro-metrics of bureaucratic quality would then either be comparable across cases by institution or suitable for aggregation to produce something like a more national-level sense of overall bureaucratic effectiveness.

State Capacity in a Large-N Dataset. These tools are, however, quite impractical for the creation of a larger dataset, especially one for which a time-series component is important. A crucial limitation on experimental techniques, thus, is that they are not suitable for historical investigation. But, of course, institutional development is generally thought of as a long-term process. Nor are they typically suitable for replication across a large number of cases. What then is an appropriate strategy? Here we think it inevitable that state capacity be measured through a focus on observable outputs of public bureaucracies that are regularly collected for a large number of countries. This could include data on educational attainment (primary and secondary enrollment, human capital stock), basic public health (infant mortality, morbidity, life expectancy), or taxation (the structure of taxation by type, the weight of taxation in the broader economy, and most importantly, the efficacy of taxation in terms of actual versus theoretical yield). But none of these data are, however, particularly useful in an untransformed form. For, taking education as an example, the ease of delivery is not constant across countries and must be understood in relationship to available material resources. Thus, our recommendation would be to pursue a metric based on the residuals from the estimation of a model of, for example, secondary enrollment rates. Here we would include as predictors of the enrollment outcome the level of economic development (to capture the resources potentially available to the state), urbanization rates (as there are scale economies in educational delivery), and a series of other controls that would get to the underlying difficulty of providing education (e.g., ethnic and linguistic fractionalization, etc.).[15] What would be of interest, then, would be the residuals from this model; that is, information on how much a particular set of state institutions over- or under-perform the expectations generated by a model based simply on resources and the difficulty of service delivery. This, we would argue, can be importantly attributed to the quality of the institutions of service delivery – in this case the educational bureaucracy. The scores on these residuals for each country and year could be normed to the overall minimum and maximum and thus placed on a relatively intuitive scale for use in subsequent analyses. As above, however, this approach is only useful in institutional contexts that are not excessively politicized, and that represent activities generally seen as core functions of most any state. It also requires clear knowledge of the non-institutional determinants of the outcome of interest if confounds are to be avoided.

[15] This would be a measure of the ability of the state to undertake its educational functions in a broad form: it is not simply a measure of the effectiveness of the educational apparatus, but also, implicitly, of the revenue-generating authorities, for without fiscal inputs education is impossible. If a very discrete measure of just the educational-providing institutions of the state is desired, it would be appropriate to control for the relevant budgetary outlays (per student) in the model, thus removing the effects of the revenue bureaucracy from the measure of the efficiency of educational provision.

Only in this way can the impact of public policy choices be at least partially mitigated –by holding it constant in comparisons– and allowing variation caused by institutional efficacy to come through. Our point, however, is not that there are easy solutions to the problem of measuring stateness, but rather that there *are* objectively-based and potentially valid (in the measurement sense) approaches that should be pursued in lieu of our current overreliance on easier, but inaccurate, and likely biased, subjective assessments.

REFERENCES

Blair, Graeme and Kosuke Imai. 2012. "Statistical Analysis of List Experiments". *Political Analysis* 20 (1): 47-77.

Corstange, Daniel. 2009. "Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT". *Political Analysis* 17 (1): 45-63.

Holsti, Kalevi. 2004. *Taming the Sovereigns: Institutional Change in International Politics*. New York: Cambridge University Press.

Huber, Evelyne. 1995. "Assessments of State Strength". In Peter H. Smith, ed., *Latin American in Comparative Perspective: New Approaches to Methods and Analysis*. Boulder, CO: Westview Press.

Kurtz, Marcus J. and Andrew Schrank. 2007. "Growth and Governance: Models, Measures, and Mechanisms". *Journal of Politics* 69 (2 [May]): 538-554.

Mann, Michael. 1993. *The Sources of Social Power, Volume II: The Rise of Classes and Nation-States, 1760-1914*. New York: Cambridge University Press.

North, Douglass, John Wallis and Barry Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded History*. New York: Cambridge University Press.

Portes, Alejandro and Lori Smith. 2008. "Institutions and Development in Latin America: A Comparative Analysis". *Studies in Comparative International Development* 43 (2): 101-128.

Przeworski, Adam. 1986. "Some Problems in the Study of Transition to Democracy". In *Transitions from Authoritarian Rule: Comparative Perspectives*. Guillermo O'Donnell, Philippe Schmitter, and Laurence Whitehead, eds. Baltimore, MD: John's Hopkins University Press.

Rauch, James and Peter Evans. 2000. "Bureaucratic structure and bureaucratic performance in less developed countries". *Journal of Public Economics* 75 (1): 49-71.

Soifer, Hillel. 2008. "State Infrastructural Power: Approaches to Conceptualization and Measurement". *Studies in Comparative International Development* 43: 231-251.

Soifer, Hillel. 2006. "Authority over Distance: Explaining Variations in State Infrastructural Power in Latin America". Unpublished Ph.D. Dissertation, Department of Government, Harvard University.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology. Volume I*. Guenther Roth and Claus Wittich, eds. Berkeley, CA: University of California Press.

Marcus Kurtz is Associate Professor at the Department of Political Science, Ohio State University.
E-mail: kurtz.61@polisci.osu.edu

Andrew Schrank is Professor at the Department of Sociology, University of New Mexico.
E-mail schrank@unm.edu