

NÚMERO 224

ANDREAS SCHEDLER

The Limits to Bureaucratic Measurement  
Observation and Judgment  
in Comparative Political Data Development

NOVIEMBRE 2010



[www.cide.edu](http://www.cide.edu)

• Las colecciones de **Documentos de Trabajo** del CIDE representan un medio para difundir los avances de la labor de investigación, y para permitir que los autores reciban comentarios antes de su publicación definitiva. Se agradecerá que los comentarios se hagan llegar directamente al (los) autor(es).

• D.R. © 2010. Centro de Investigación y Docencia Económicas, carretera México-Toluca 3655 (km. 16.5), Lomas de Santa Fe, 01210, México, D.F.  
Fax: 5727•9800 ext. 6314  
Correo electrónico: [publicaciones@cide.edu](mailto:publicaciones@cide.edu)  
[www.cide.edu](http://www.cide.edu)

• Producción a cargo del (los) autor(es), por lo que tanto el contenido así como el estilo y la redacción son su responsabilidad.

---

*Acknowledgements*

*A previous version of this paper was prepared for delivery at the 106<sup>th</sup> Annual Meeting of the American Political Science Association (APSA), 2-5 September 2010, Washington, DC. I thank Alejandro Anaya, Carlos Bravo, Ron Francisco, Jennifer Gandhi, Carolina Garriga, James Melton, Covadonga Meseguer, Alvaro Morcillo and Gilles Serra for most valuable comments. The usual disclaimers apply.*



## Abstract

---

*Standard methodological advice in the social sciences tells us to base our measurement decisions on observations, not judgments. It presupposes that we divide the process of measurement into two phases. In a first judgmental stage, we reach all the judgments necessary to build our theories, form our concepts, delineate our objects of observation and devise our coding rules. In a second observational stage, by contrast, we suspend our judgmental faculties in order to apply our formal rules of measurement in a quasi-bureaucratic fashion. However, does the full bureaucratic regulation of observation and measurement represent a feasible ideal in the comparative study of politics? In the present paper, I explore the limits to "pure" observation in three steps. In the first section, I describe the "bureaucratic ideal" of social measurement. In the second section, I review four conditions that may compel scholars to incorporate elements of judgment into their measurement processes: unobservable realities, unobserved realities, unexpected realities and complex concepts. In the third section, I offer some thoughts about methodological standards that may guide judgmental modes of comparative measurement.*

## Resumen

---

*De acuerdo a estándares metodológicos comunes en las ciencias sociales, debemos elaborar nuestras mediciones exclusivamente con base en observaciones empíricas, prescindiendo de todo elemento de juicio personal. El presente ensayo describe este ideal metodológico como "burocrático", dado que aspira a eliminar la discreción humana por medio de reglas formales. ¿Hasta qué punto se trata de un ideal viable que pueda guiar la generación de datos comparados de manera realista? El documento explora cuatro factores que imponen límites, sean intrínsecos o pragmáticos, a la medición burocrática: realidades no observables, realidades no observadas, realidades inesperadas y conceptos complejos. Las cuatro condiciones nos pueden llevar a incorporar el juicio de expertos en el desarrollo de datos. La última sección del documento ofrece algunas reflexiones sobre estándares metodológicos que podrían guiar procesos de medición basados en el juicio de expertos.*



## *Introduction*

---

Standard methodological advice in the social sciences tells us to base our measurement decisions on “observations, rather than judgments” (Przeworski *et al.*, 2000: 55). It presupposes that we divide the process of measurement into two phases. In a first judgmental stage, we reach all the judgments necessary to build our theories, form our concepts, delineate our objects of observation, select the empirical phenomena we admit as observational evidence and devise the coding rules that allow us to assign numbers to cases. In a second observational stage, by contrast, we suspend our judgmental faculties in order to apply our formal rules of codification in a mechanical fashion. In the present paper, I describe this methodological ideal as “bureaucratic” because it embraces the core aspiration of bureaucratic regulation: eliminating human discretion through the formal regulation of decision-making. Reframing and rethinking our measurement standards as “bureaucratic” standards may seem no more than a rhetorical move. My intention is not polemical, however. I believe that the deep-seated analogies between our methodological ideals of data development and our bureaucratic ideals of human decision making help us to better understand the logic of social scientific measurement, its ambitions and achievements as well as its limitations.

The main issue I wish to address in this paper concerns the intrinsic and practical limits to “pure” rule-based observation. Does the full bureaucratic regulation of observation and measurement represent a feasible ideal in the comparative study of politics? Under which conditions may judgment (understood as informed and reasoned decision-making) come to represent a necessary (and even desirable) part of comparative political measurement? I shall explore the limits to nonjudgmental observation in three steps. In the first section, I describe the “bureaucratic ideal” of social measurement. In the second section, I review four conditions that may compel scholars to incorporate elements of judgment into their measurement processes: unobservable realities, unobserved realities, unexpected realities and complex concepts.

In the third section, I offer some thoughts about methodological standards that may guide judgmental modes of comparative measurement. Mind that I do not introduce a methodological alternative to “bureaucratic” measurement. Almost to the contrary, after specifying the conditions that may require us to turn to judgmental modes of data development, I propose to incorporate bureaucratic standards into the construction of judgmental data. In this sense, the paper yields an ironic conclusion: When the limits of bureaucratic observation compel us to rely on judgmental measurement, we

should think about possible strategies of bureaucratizing the judgment of experts.

### ***1. The Bureaucratic Ideal of Measurement***

In the world of physical objects and standardized units of measurement, measurement is traditionally understood as “the practice of attempting to identify the magnitude of a quantitative attribute by estimating the ratio between that magnitude and an appropriate unit” (Michell, 2005: 678). In the social sciences, we lack standardized units of measurement for most purposes of quantification. Commonly, we cannot tell how such units of measurement might even be conceived in the first place. Thus, for the enterprise of social measurement, the wider, classic definition offered by psychologist Stanley S. Stevens seems more appropriate: “Measurement is the assignment of numbers to objects or events according to rules.”<sup>1</sup>

In both physical and social measurement, numbers are supposed to reflect properties of objects. They are not to be determined by the vicissitudes of the measurement process. Accurate measurement establishes a relation of correspondence between the symbolic realm of arithmetic numbers and the empirical realm of objective attributes. It requires reliability, that is, the reduction of measurement error (inconsistency between numbers and object properties) induced by either faulty measurement instruments or their faulty application.

In the realm of physical measurement, reliability demands technical instruments competent agents can apply with low margins of error. In the realm of social measurement, it demands explicit formal rules competent agents can apply with low margins of error. Reliable social measures are impersonal measures. They do not vary with the identity of the person who performs the act of measuring. Reliable procedures of social measurement produce (almost) identical results (with low random error) when different people assign numbers to identical observations on the basis of identical rules of measurement.

The methodological ideal of social measurement is analogous to the political ideal of bureaucratic decision making. Just as the ideal bureaucracy involves a perfect separation between politics and administration, the ideal measurement involves a perfect separation between judgment and observation. According to the idealized conception of bureaucracy, politicians define the rules, while lower-level public officials only apply them in a mechanical, impartial and impersonal manner. Politics is the realm of discretionary decision making, administration the realm of non-discretionary

---

<sup>1</sup> Cited in Neuendorf (2002: 111). Note that Stevens’ rule-based rather than unit-based definition does not distinguish measurement from counting (the determination of frequencies of empirical phenomena that are identified as members of the same conceptual category).



rule application. According to the idealized conception of social measurement, empirical researchers define the rules, while research assistants (measurers) only apply them in a mechanical, impartial and impersonal manner. Theory building and concept formation is the realm of discretionary decision making, measurement the realm of non-discretionary rule application.

The notion that scientific judgment and bureaucratic observation can be held apart in a neat manner rests upon a series of presuppositions, both about the nature of reality and the nature of our concepts. First of all, it presupposes that our concepts form *classical typologies*: complete and consistent hierarchies of mutually exclusive categories that are neatly ordered along a “ladder of abstraction” (Giovanni Sartori), with the abstract basic concept at its top, its concrete indicators at its base and as many intermediate levels of abstraction in between as deemed necessary. Figure 1 (in the appendix) shows a simple conceptual hierarchy with one intermediate level of constitutive dimensions, each branching out into a pair of empirical indicators. The conceptual map is clear and complete. Conceptual dimensions are comprehensive and mutually exclusive and their empirical indicators too. Abstract dimensions are clearly anchored in concrete observations; and the other way round, concrete observations are clearly assigned to abstract dimensions. There are no overlaps and no duplications, no free-floating categories and no homeless observations.<sup>2</sup>

Well-ordered conceptual hierarchies allow us to anchor abstract concepts in concrete realities (several steps down the ladder of generality) and the other way round, to tie concrete observations to conceptual abstractions (several steps up the ladder of generality). As a matter of course, constructing or reconstructing consistent classificatory schemes is not a bureaucratic enterprise, but a creative one that requires reflection and judgment (see Schedler, 2011). Yet if we wish to apply our conceptual schemes to empirical realities and quantify them in a manner that relies on “pure observation” and *precludes* further creativity, reflection and judgment, four additional conditions must hold.

For political measurement to operate in a non-judgmental fashion, we need (1) transparent empirical phenomena whose observation does not depend on our judgmental faculties; (2) complete public records on the empirical phenomena we wish to measure; (3) complete, consistent and determinative rules of measurement that eliminate discretion in the assignment of numbers to the phenomena we observe; and (4) low levels of complexity (abstraction, dimensionality and aggregation) of our concepts. In the comparative study of politics, these conditions are often problematic. Let

---

<sup>2</sup> Prominent examples of work on conceptualization and measurement that rest upon these assumptions are Sartori (2009a) and (2009b), Goertz (2006) and Munck and Verkuilen (2002).

us briefly review them one by one (for an overview, see Table 1 in the appendix).

## ***2. The Judgmental Demands of Measurement***

If social measurement is *the assignment of numbers to observations according to rules*, the methodological imperative according to which scientific measurement requires *observation, not judgment* is *incomplete*. Measurement requires the ability to observe (a faculty of the eye), but also the ability to follow rules (a symbolic competence). If judgment involves the exercise of discretion, we can only dispense of it in either of these *two* operations (observation and rule application) if facts are obvious and rules determinate. Otherwise the application of rules is not “mechanical”, but requires elements of independent judgment. This is true for the production of measurement decisions as well as for administrative or judicial decision-making. Yet, in the realm of social measurement, just as in bureaucratic organizations and judicial systems, the ideal conditions for rendering human decision making irrelevant (non-decisive) often do not hold empirically (see also overview Table 1).

### *2.1. Unobservable Realities*

The primacy of observation is the hallmark of methodological positivism. What we see is what we believe in (see also Johnson, 2006). Privileging the eye over other human organs may be a plausible methodological choice —as long as there is anything *to see*. In the study of politics, though, most of the empirical phenomena we are interested in are not accessible to direct observation. Not to “pure” observation anyway.

#### *a. Meaning*

Social sciences are exercises in “double hermeneutics” (Anthony Giddens): we try to make sense of the others (and ourselves) who are trying to make sense of others (and themselves). We do not study inanimate, objectively given realities, but symbolic, socially constructed realities. To grasp them, we need to understand them. Simple and pure observation won’t do. Men from Mars can’t practice social sciences. Not on earth, anyway. They need more than to register the outward movements of men and objects in order to understand the actions and institutions they are watching. For instance, by merely witnessing multitudes of people marking sheets of paper and dropping them into transparent boxes, they cannot comprehend the institution of political elections. Whenever we talk of “observation” of social or political realities, it is never “pure” observation we are referring to, but meaningful observation,

guided by our pre-existing social knowledge and conceptual tools. In this sense, we are all interpretivists. We all interpret pre-interpreted realities. This may be a trivial thing to state.<sup>3</sup> We all know that facts do not speak for themselves. Yet, since we sometimes speak as if they would, it seems pertinent to recall the intrinsically interpretative nature of our observations.

### *b. Causes and Counterfactuals*

Notoriously, we cannot observe causation nor the counterfactual worlds causal reasoning involves (such as the absence of effects in the absence of necessary causes). Our disciplinary reflections on “causal inference” accordingly revolve around problems of explanation under conditions of limited observation (see King, Keohane and Verba 1994). Yet, the non-observable nature of causal relations and counterfactual worlds creates methodological challenges not only at the level of explanation, but also at the level of description (and thus measurement). Causes and counterfactual conditions are often integrative parts of the phenomena we try to describe and explain. They are built into the very concepts we try to measure.

For example, the concept of vote buying assumes that buyers and sellers of votes establish effective relationships of commercial exchange (see Schaffer and Schedler, 2007). If we wish to measure “vote buying” it is not enough to estimate either the amount of clientelist investments parties and candidates realize or the magnitude of electoral support they receive. We need to establish the causal relationship between the two (which is a demanding enterprise that involves, among other things, counterfactual reasoning about voter choices in the absence of vote-buying efforts). Any effort to measure concepts that rest upon causal assumptions (be they overt or hidden) cannot live on observation alone.

### *c. Subjectivity*

The bounded world of subjective beliefs, values and emotions is shut off from external inspection. There are two main indirect ways for gaining access to the realm of subjectivity. We can take seriously what people tell us about their thoughts, desires and feelings and we can ask them and listen. Or we can search for outward symptoms of their inner states and look what they do and how they behave. Yet, notoriously, regardless of how we try to comprehend what goes on inside the minds and hearts of others, we cannot directly *observe* it. We can only *infer* it from what we hear and see.

Scholars who investigate the realm of subjectivity, as in the study of public opinion, are familiar with the difficulties of making visible phenomena that are essentially invisible. The non-observable nature of research objects is

---

<sup>3</sup> It am aware that it may also be a polemical thing to state. Yet I am bracketing here the rather arcane methodological debates on the notion of interpretation that define imaginary cleavages in us political science.

less obvious in those cases in which ostensibly observable phenomena contain inbuilt elements of subjectivity. For instance, the notion of political violence refers to observable acts as well as to subjective motivations. Its political motives (however defined) distinguish political violence from other forms of violence, such as domestic violence, criminal self-enrichment through the use of force, or irrational acts of violence without discernible motives. What we can *see* (although in most instances, luckily, not in the first person) is the exercise of violence. What we need to *infer* (often based on knowledgeable judgments by others) are the motives that drive the exercise of violence.<sup>4</sup>

## 2.2. Unobserved Realities

Some empirical phenomena we cannot observe in principle, others we cannot observe in practice. More often than not, the high informational demands of bureaucratic observation cannot be met in reality. In particular, in the comparative study of politics, empirical information is often incomplete or inconsistent.

### *a. Incomplete Information*

Despite the dizzying expansion of cross-national political datasets in the past years, we lack basic information on innumerable questions in comparative political inquiry. Entire spheres of politics and categories of data are off our screens. For instance, we suffer from chronic and systematic information shortages with respect to *(i)* political phenomena that are hidden from the public view due to their illicit nature, such as crime and corruption,<sup>5</sup> *(ii)* political phenomena, such as contentious actions or subnational processes, that are observed by domestic private agencies, but hard to establish from the outside by international researchers and *(iii)* official data on state institutions and decisions, such as judiciaries and judicial outputs, that are generated by national public agencies, but not pooled at the international level (see Schedler, 2012).

If no data are available on any case, researchers cannot but resign themselves to ignorance. In the complete absence of information on political phenomena, be it direct or indirect, they must renounce the pretension of measuring them. Alas, usually we do have at least some bits and pieces of

---

<sup>4</sup> Note, though, that the contemporary debate about private versus political explanations of civil war (“greed” versus “grievance”) (see Collier and Hoeffler, 2004) tends to obliterate the distinction between private and political violence.

<sup>5</sup> Of course, most illicit acts do not go “unobserved” in a literal sense. Someone is watching and be it just the perpetrator and the victim. Yet, private observations do not count in scientific measurement. As any other bureaucratic enterprise, scientific observation requires the written form. As Max Weber had it, the public record (Schriftlichkeit) constitutes a key operating principle of modern bureaucracies. When we talk of scientific “observations” we talk of recorded and publicly accessible information.

information on at least some cases we are interested in. In such situations of incomplete information, we have to bridge informational gaps either by relying on contextual knowledge (expert judgment), or by devising general rules that deal with incomplete information (rules of ignorance). The creation of ad hoc rules in response to emergent measurement problems is a matter of judgment. Yet, once created, those rules can be used to resolve analogous problems in the future (as well as, if necessary, to recode data collected in the past).

*b. Inconsistent Information*

Data from different sources are likely to diverge due to observational error. Moreover, in the study of politics, they are likely to diverge due to political bias. The providers of information are often parties to political struggles in which information itself represents an essential resource. For example, social movements and government agencies notoriously tend to diverge in their estimates of attendance to anti-government demonstrations. In such situations of inconsistent information, once again researchers can arbitrate between diverging accounts either by relying on contextual knowledge (expert judgment), or by devising general rules that deal with informational inconsistencies (rules of arbitration).

*2.3. Fuzzy Boundaries*

In a complex and changing world, we should always be prepared to encounter cases that do not quite fit our conceptual boxes and operational guidelines. We create our concepts and trace their boundaries, but we do not create and control the realities we try to measure. It's like the settlement of legal disputes: universal law can never foresee the infinite variation in particular cases that may arise in the future. Thus the need for judicial decision-making. Our universal rules of measurement try to grasp heterogeneous and evolving realities. They may fit well the standard situations we had in mind in devising them, but we are likely to encounter problems of application in the "gray zone" (Goertz, 2006) of non-standard cases in which our generalizing assumptions about the structure of the world do not hold. Even ostensibly simple tasks of establishing the presence or absence of a phenomenon may turn problematic. For instance, when counting opposition parties in non-democratic regimes we may encounter parties that belong to the opposition in name, yet are created, sponsored, directed, subverted, or manipulated by state agencies (see Wilson, 2005). Again, we can make sense of such "hard" or borderline cases either by relying on our contextual knowledge (expert judgment), or by devising general rules that take care of problems of operational delimitation (boundary rules).

As stated above, the observation of social realities depends on our interpretative faculties, our capacity to understand symbolic realities. The elementary dependence of observation on interpretation cannot be bridged, circumvented, or mitigated by any amount of bureaucratic regulation. Whether we like it or not, whether we recognize it or not, we have to live with it. There is no escape either from the fact that we cannot observe causal relations or counterfactual worlds. Establishing causes and counterfactuals is a matter of argumentation (causal inference), not observation. By contrast, all other challenges to bureaucratic rule application in political measurement that I have discussed so far, in principle do know bureaucratic remedies.

We cannot directly observe subjective realities, but we can often devise indirect indicators, observable (visible and readable) symptoms that reveal the nature of underlying subjective realities. We often have to deal with incomplete or inconsistent information and cannot foresee all possibilities of informational gaps and contradictions. Yet we can devise rules of ignorance and rules of adjudication that anticipate some of these informational problems; and we can amend these rules if we encounter novel problems in the process of measurement, so that we have general rules to guide us when we (or others) stumble over similar problems at another time. The same applies to hard cases that inhabit the gray zone between categories or measurement scores: If we encounter borderline cases that are difficult to make sense of on the basis of existing coding rules, we can amend these rules and thus provide formal and explicit guidance for similar measurement decisions in the future.

These regulatory bridging devices allow us to measure what we cannot observe, and to create supplementary rules of measurement when the application of existing rules runs into difficulties. They work well as long as levels of conceptual and empirical complexity remain moderately low. Yet, they turn unfeasible in the face of complex concepts that are abstract, multi-dimensional and aggregate.

#### *2.4. Complex Concepts*

In scientific measurement, just as in real life, we can try to eliminate discretion and surprise by weaving dense webs of authoritative regulation. As long as the phenomena we strive to subject to formal and explicit rules are relatively simple, stable and neatly bound, regulation may indeed work as an effective device to create predictability and constrain human agency. In politics, at higher levels of complexity, the legislative pretension of foreseeing everything and regulating everything is certain to create bureaucratic nightmares. Recall the tragic comedy of real-socialist command economies. In an analogous fashion, when the concepts we wish to measure reach a certain degree of complexity, the methodological pretension of

devising a full catalogue of coding rules that establishes clear and precise links between all possible elements of empirical evidence (including all possible gaps and inconsistencies of evidence) is likely to produce a bureaucratic nightmare, too. In the face of complex realities, the notion of complete, consistent and determinate law represents an idealized fiction—in the realm of legal regulation as much as in the realm of methodological regulation.

*a. Conceptual Complexity*

In comparative political measurement, simple concepts can be quantified in bureaucratic fashion. They can be measured through the mechanical application of formal rules that relate observations to numbers without giving rise to doubt or ambiguity. As a matter of fact, though, it is not easy to come up with clear-cut examples. Nevertheless, the holding of elections, the abolition of national legislatures, and the declaration of international war (all categorical events) may count as plausible exemplifications. The measurement of complex concepts, by contrast, is generally not susceptible to full bureaucratization (except at the price of radical reductions of complexity). The observation and measurement of complex concepts imposes informational demands that cannot be processed through formal regulation. They can only be met by expert judgment, grounded in analytical and synthetic competence as well as in local knowledge.

Conceptual complexity entails three structural properties: abstraction, composition and aggregation.

- A. *Abstraction*: Complex concepts are abstract, rather than concrete. Situated at high levels of generality, they oblige researchers to travel a long way on the road from conceptualization to operationalization. To get from the general concept to concrete indicators they have to laboriously descend the ladder of abstraction by multiple steps.
- B. *Multi-dimensionality*: Complex concepts involve multiple, rather than single, dimensions (and subdimensions). The challenge of anchoring abstract ideas in concrete realities multiplies by the number of dimensions (and subdimensions) a concept accommodates.
- C. *Aggregation*: Complex concepts refer to aggregate, rather than singular, phenomena. Aggregation may be spatial (across territories), temporal (across time), or social (across groups of actors). In the comparative study of politics, concepts routinely refer to properties of national political systems. These macro-level properties often do not capture single events at the center stage of national politics. Rather they represent aggregate results of countless events that take place in a decentralized fashion, on countless locations far off the capital city.

Measurement thus involves the challenge of collecting and aggregating streams of information that tend to be overwhelmingly rich and despairingly incomplete at the same time.

These three structural features of conceptual complexity may vary independently of each other, but frequently they do go together. Consider the example of *electoral quality*. Within the liberal consensus, the existence of “free and fair” elections represents a constitutive dimension of modern representative democracy. Political elections are notoriously complex processes. The expanding archive of case reports produced by the community of election observers over the past two decades as well as the expanding scholarly literature on electoral manipulation have borne testimony to the complexity of empirically ascertaining the democratic quality of elections - in particular in the vast border area between clearly democratic and clearly dictatorial elections.

For instance, election experts Jørgen Elklit and Andrew Reynolds affirm, quite persuasively, that the “systematic study of election quality” requires the systematic study of eleven “steps” or dimensions of electoral processes: legal framework, electoral management, constituency and polling district demarcation, voter education, voter registration, ballot design, party and candidate nomination and registration, campaign regulation, polling, vote counting and vote tabulation, electoral dispute resolution and post-election procedures. According to the authors, to ascertain the democratic quality of these dimensions we have to address a total of fifty questions or “performance indicators” (a bit less than five per dimension). Many of them, such as questions they include about levels of violence and voter intimidation, contain further subdimensions and demand collecting information on the entire voting process, with thousands of locations and millions of actors (see Elklit and Reynolds, 2005: Table 1).

The complexity of the measurement task at hand is staggering. If we would try to devise coding rules that would allow election observers to assign numbers to all 50 performance indicators in a mechanical fashion, without exercising discretionary decision-making power, we would need at least a dozen of rules for each indicator. Most likely, we would need many more. If it were only ten rules per indicator, we would need 500 coding rules (!) to assess the democratic quality of an election (and then, of course, some more to determine their aggregation). A regulatory and administrative nightmare, a recipe for mental as well as methodological insanity. No wonder that Elklit and Reynolds reach the conclusion that most of their indicators have to be ascertained through “expert panel assessments” (*Ibid.*).



*b. Bureaucratic Shortcuts*

Are there any methodological shortcuts that allow us to measure complex concepts bureaucratically, without recurring to the judgmental faculties of experts? There are some, indeed, although they come at the cost of radical simplification.

- A. *Conceptual jumping*: We may bridge the complexities of abstraction through “conceptual jumping,” that is, by drawing direct linkages between abstract concepts and concrete indicators, while ignoring all intermediate levels. Although the most abstract or “basic level” of conceptualization is usually “too abstract and complex to be directly converted” into specific indicators (Goertz, 2006: 53), conceptual jumping has been the “dominant” approach in statistical research where it is habitual to see scholars connect abstract concepts and quantitative indicators without mediating steps (*Ibid.*: 55).
- B. *Litmus tests*: We can circumvent the complexities of multidimensional concepts if (and only if) we can come up with reliable “litmus tests,” that is, if we can conceive of specific symptoms whose presence firmly indicate the presence of the general concept we wish to measure. If we can observe such symptoms, we need not be able to observe the underlying condition that produces them. Here, in the compelling logic of symptoms, resides the ingenuity of Adam Przeworski and his collaborators’ (2000) designation of alternation in executive power as key indicator of the democratic quality of elections. Authoritarian elections preclude opposition victories at the polls (substantive certainty), democratic elections make them possible (substantive uncertainty).
- C. *Samples*: We may reduce the complexities of aggregation by devising “rules of representation,” that is, rules and procedures that allow us to select a subset of observations we can plausibly take as “representative” of the whole. Random sampling, the selection of worst observations, and the selection of end-of-year observations are examples of such strategies.

Circumnavigating the complexity of concepts through bureaucratizing remedies (conceptual jumping, litmus tests and sampling strategies) involves radical reductions in conceptual and operational complexity. If we are not prepared to accept the courageous simplifications (and ensuing losses of validity) they often impose, if we wish to measure complex concepts—such as state capacity, the rule of law, electoral integrity, civil society and many others— at higher levels of complexity, we must rely on the local knowledge

and judgmental faculties of experts. Of course, judgmental data have a terrible press, in particular among quantitative methodologists who tend to describe (and disqualify) them as “subjective” (Bollen and Paxton, 2000) and to issue urgent calls for “bringing objectivity back in” (Kurtz and Schrank, 2008: 8). The question is whether can we reach measurement judgments in valid and reliable ways, without giving reign to subjective arbitrariness.

### ***3. The Reliability of Judgmental Measurement***

According to standard measurement theory, measurement procedures are considered “reliable” to the extent that the numbers they assign to empirical objects correspond to “real” attributes of these objects. For example, a history exam is reliable to the extent that variations in the grades different students obtain correspond to “true” differences in their historical knowledge, rather than being artifacts of the examination procedure itself. Yet, as “true” scores of anything are generally unknown, scholars generate estimates of reliability by using repeat tests. Standard “reliability coefficients” do not report the degree of convergence between measures and realities, but between measures obtained in iterated measurement efforts (commonly carried out by different researchers). In practical terms, the notion of reliability thus overlaps with the idea of replicability. What we may thus call the methodological norm of “reliable replicability” demands, firstly, that our measurement procedures are transparent enough to allow others to repeat them and, secondly, that the results of replication lie close to the original results (with low margins of random error).

Reliable replicability constitutes the core value (and main achievement) of bureaucratic measurement procedures. The bureaucratization of political measurement does not guarantee its validity. It does ensure, though, that different researchers who measure the same phenomenon on the basis of shared conceptual choices, empirical evidence and measurement procedures are likely to reach similar results. The central value of judgmental measurement, by contrast, is validity—the validity of informed and reasoned public argument. Expert judgments do not strive to be replicable, but persuasive. Nevertheless, just as bureaucratic measures cannot ignore question of validity, judgmental measures do have to address issues of reliability too. Judgmental measurement is of little use if the numbers experts assign to cases correspond to patterns of scientific judgment, rather than patterns of empirical reality. In the face of complex concepts and complex realities, the general case for judgmental measurement is strong. Still, methodologists will continue to raise plausible objections to the “subjective” nature of expert judgments, unless we succeed in “disciplining” expert judgment and hereby increasing confidence in the quality of experts as well as in the quality of their measurement decisions.

In bureaucratic measurement, we obtain reliable measures to the extent that (i) we hold observations (or procedures for sampling observations) constant across repeated acts of measurement, (ii) our coding rules are complete, precise and consistent and therefore permit the mechanical (non-discretionary) assignment of numbers to observations and (iii) our coders are competent observers and rule followers.<sup>6</sup> These standards basically hold for “subjective” measures as well, that is, for data that collect personal perceptions individuals drawn from variously defined non-expert groups carry with respect to political phenomena. Sometimes, such “subjective” data aggregate judgments lay persons hold with respect to complex political realities, but they do not issue judgments themselves. They only shift the burden of judgment from the researcher to the general public or segments of it.<sup>7</sup>

If we aspire to subject judgmental measurements to standards of reliability, it would be a categorical mistake to apply the same standards we apply to bureaucratic measurement. If we start thinking about plausible reliability standards for judgmental measurement, we have to start by recognizing that these standards have to be different from those that govern the bureaucratic collection of data. Expert judgments are often confused with “subjective” data. Yet, in contrast to subjective measures, they are not supposed to be subjective, but intersubjective: grounded on public facts and public reasons, defensible in the face of critique. Furthermore, in contradistinction to bureaucratic measures, expert judgments are not supposed to be impersonal. Coders of factual observations are fungible, experts are not. While the identity of the former must not matter for the results of factual measurement, the identity of the latter is constitutive for the construction of judgmental data.

Proposing a full set of procedural guidelines for judgmental measurement lies beyond the scope of the present paper. Still, I would like to indicate some obvious areas for methodological improvement.

---

<sup>6</sup> These conditions are quite demanding and often are not met in practice (see Lieberman, 2010; Pepinski, 2007; Schedler, 2012).

<sup>7</sup> The reliability of subjective data is not problematic per se. It simply depends on the controlled nature (context-independence) of the accompanying procedures and the representative nature of the underlying sample. Unless procedures are constant and unobtrusive, individual responses will vary every time we ask. Unless samples are representative, aggregate measures will vary every time we collect them. For decades, survey researchers have been refining their sampling techniques, survey design and interviewing procedures. Cross-national surveys like the World Values Surveys and the Global Barometer surveys have been striving hard to ensure their international partners follow recognized methodological standards. To the extent that their technical foundations have been transparently solid, the reliability of cross-national political surveys has been a much lesser concern than their validity and comparability. For overviews on cross-national public opinion research, see Heath, Fisher and Smith (2005) and the symposium on “The Proliferation of Comparative Survey Research” ASPA-CP Newsletter 15/2 (Summer 2004). 5-25.

*a. Expert Selection*

The quality of expert judgments depends, first of all, on the quality of experts. The identification and selection of genuine experts is key to the production of genuine expertise. Expert polls as well as other forms of judgmental measurement often fail on this account. For instance, while the notion of representative random sampling makes limited sense in expert communities (either you know or you don't), expert polls face recurrent critiques against their use of "very small samples of individuals" (Landman and Häusermann, 2003: 30). In response, following the misguided notion that "more is better," they sometimes strive to maximize their number of observations, which makes them prone to include individuals with highly varying degrees of knowledge on the specific theme under investigation. The generous notion of "country expertise" often suffices to incorporate individuals into expert surveys that demand deep knowledge on specific substantive issues *within* countries. In particular in small and poor countries with fragile social scientific infrastructures, only very few individuals, if any, are likely to command the requisite information on specific policy fields or institutional arenas.

In addition to improving *ex ante* mechanisms of expert selection, we may also improve *ex post* mechanisms of expert accountability. Guarding the guardians, judging the judges, might be a possible (though complicated) response to manifest variations in knowledge (as well as impartiality) of experts. As current APSA president Henry Brady once suggested, when commissioning country "ratings from knowledgeable observers... we might want to think about whether we should scale the raters as well as the countries that are rated" (2004: 64).

*b. Comparative Anchors*

For expert judgments to serve the purpose of cross-national comparison (or any kind of comparison, for that matter), they must be, well, comparable. That is, the numbers they assign to countries or phenomena within countries must make sense across countries (or other units of observation). If the same numbers mean different things to experts from different countries (or to different experts within one country), they are useless for comparative purposes. For example, if we ask reputed experts to estimate whether the incidence of vote buying and voter intimidation in a national election campaign was "high", "moderate", or "low", we need shared standards (in addition to shared concepts) to ensure that their judgments are meaningful in comparative terms. Observers may judge prevailing levels of electoral clientelism and intimidation on the basis of various criteria, such as democratic ideals, past experiences, paradigmatic cases, or regional averages. In the absence of normative anchors, of common, explicit and transparent benchmarks, their judgmental measures will be as "subjective" as

their critics suspect them to be. They will lack the intersubjective quality that defines meaningful judgmental measurement.

*c. Transparency*

Reliability is a standard we demand from the repeated application of measurement procedures to invariant empirical phenomena. If we apply identical procedures to varying phenomena, we have no reason to expect the results to be reliably similar. Cross-national political datasets often fail to disclose their information sources in a systematic and transparent manner. Frequently we learn that dataset authors rely upon a certain range of information sources, without getting to know the precise information bases that motivated specific coding decisions. We learn about the rough contours of their camp of vision, but cannot know what exactly they have been looking at when taking concrete coding decisions. In the end, we cannot relate numbers to observations in a precise fashion.

The difficulties of relating sources and scores tend to be rather obvious in the case of judgmental measures. For example, in its more recent annual reports on the state of freedom in the world, Freedom House publishes selective listings of more than 200 periodical publications and over 120 organizations that go into its global estimates of political rights and civil liberties (see *e.g.* Piano, Puddington and Rosenberg, 2006: 902-906). Of course, one of the key assets experts bring into the measurement process lies precisely in their capacity to process and synthesize large amounts of dispersed information. At the end of the day of measurement judgment, it is usually impossible for them to relate the numerical conclusions they reach to the precise pieces and bits of information that have gone into them. Nor are they in a position to provide an algorithm that would trace the mental process of reasoning that led them from the assessment of empirical evidence to the assignment of scores. And yet. Even if experts are unable to describe all the miniature pieces that comprise a complex mosaic of knowledge generation and analytic judgment, they should be able to document the big picture. They should be able to provide, not *all* sources and modes of reasoning that have gone into their measurement decisions, but the central ones. And just like historians, they should be able to explain the range of uncertainty and controversy of their judgmental decisions with reference to concrete documentary evidence (or the lack of such evidence).<sup>8</sup>

---

<sup>8</sup> For a systematic treatment of proximity and transparency of data sources, see Lieberman (2010).

#### *d. Aggregation*

Since experts may not fully converge in their assessments, data producers must have some way of aggregating their diverging judgments. Standard textbook advice of adjudicating among diverging coding decisions through random procedures seems reasonable in the case of bureaucratic measurement, yet makes little sense in the case of expert judgment.<sup>9</sup> Expert surveys, in which data producers collect judgments by *external* experts, tend to rely on *additive* procedures (the calculation of arithmetic means) that grant equal weight to all individual judgments. Examples are the Legislative Power Index assembled by Steven Fish and Matthew Kroenig (Fish, 2006) and the data on subnational regimes in Argentina constructed by Carlos Gervasoni (2008).

Expert studies, in which data are generated by *internal* personnel within the responsible data agency (be it research group, university department, non-governmental organization, government agency, or international organization), tend to rely on *deliberative* procedures (the reconciliation of discrepancies through communication) in which it is “the forceless force of the better argument” (Habermas, 1981: 47) that carries the final measurement decision. Freedom House scores of political rights and civil liberties ([www.freedomhouse.org](http://www.freedomhouse.org)) and the Bertelsmann Transformation Index ([www.bertelsmann-transformation-index.de](http://www.bertelsmann-transformation-index.de)) exemplify judgmental data that arise from layers of expert deliberation.

Deliberation actually seems to be the most appropriate procedure to settle disagreements among experts. However, we still need to devise procedural guidelines that render deliberation compatible with our methodological demands of transparency. Confidential deliberation by chain smoking experts in dark chambers is not enough. The quality of expert judgments is defined by the quality of their public justification. Even if deliberative processes take place within closed circles of experts, their results must obey the principle of publicity. Argumentative opacity destroys the legitimating core of expert judgments. In the end, we cannot trust the experts (blindly), but need to trust their arguments (seemingly).

---

<sup>9</sup> Random choice of diverging coder decisions is a standard procedure in other disciplines, like psychology and media research, yet virtually unknown in comparative political science. In comparative politics, divergences in measurement decisions are commonly settled either through deliberative processes or through authoritative decision-making (by the lead researcher).

### *3.1. Code: Olympic Bureaucracies*

To improve the quality of judgmental measurement in comparative politics we might venture some glances beyond our disciplinary boundaries to see how others construct their judgmental data. For instance, we might learn from the way “judgmental” disciplines of sports set up their juries. To ensure high-quality measurement of complex performances, olympic gymnastics, for instance, employs a sophisticated system of checks and balances that includes:

- Stringent procedures and requirements for the selection of knowledgeable, experienced and proven judges;
- Stringent provisions for continual training, examination and surveillance of judges, with peer committees assessing their accuracy, integrity, consistency and discipline at each high-level competition;
- A stringent codes of ethics and discipline to sanction transgressions;
- A minute definition of performance standards, evaluative concepts, rewards for original and demanding exercises and penalties for performance failures, yielding breathtakingly precise and detailed coding rules;
- Multiple teams of coders with common information among judges (through their joint presence at the competition as well as the recording of performances and their display on computer screens);
- The elimination of outliers (highest and lowest scores) as well as additive coding across teams (by averaging results) and
- The public nature of individual votes (roll call voting), with the known, disciplining presence of critical audiences possessing access to identical observations.<sup>10</sup>

In other words, even while recognizing the necessity of judgment, the international gymnastics authorities strive to bureaucratize the assignment of numbers to artistic performances as much as possible and thus to eliminate human judgment as much as possible. The passion for bureaucratic regulation is understandable, given the stakes of Olympic competition as well as the history of suspicions (and evidence) of national bias (see Sala *et al.*, 2007). In those areas of comparative political measurement where judgment seems unavoidable (or even desirable), we may not go that far. In particular, if we wish to open up a legitimate role for deliberation in data production, we might learn from possible contrast as much as from eventual analogies to judgmental sports.

---

<sup>10</sup> The website of the Fédération Internationale de Gymnastique (FIG) offers a broad collection of rules ([www.fig-gymnastics.com](http://www.fig-gymnastics.com)).

## *Conclusions*

---

With certain regularity, the modern project of forging “iron cages of bureaucracy” (Max Weber), of imposing formal rules to constrain human agency and assimilate it to the smooth operation of machines, runs into certain obstacles. Multiple obstacles, as a matter of fact. Some of them arise from the conceptual and epistemic complexities of rule application. Understanding and relating concepts, rules and realities often requires elements of judgment. In all known realms of bureaucratic regulation, the notion of self-applying rules has revealed itself as an idealized fiction, as a regulative ideal that we may approximate, yet never fully accomplish. We may deplore when public officials exercise judgment and accuse them of usurping political functions (“bureaucratic discretion”). We may deplore when judges exercise judgment and accuse them of usurping legislative functions (“judicial activism”). We may deplore when hunters and gathers of political data exercise judgment and accuse them of subjectivity (“methodological activism”). At the end of the office day, however, we will have to recognize that formal rules may serve to constrain, yet hardly to eliminate, human decision-making.

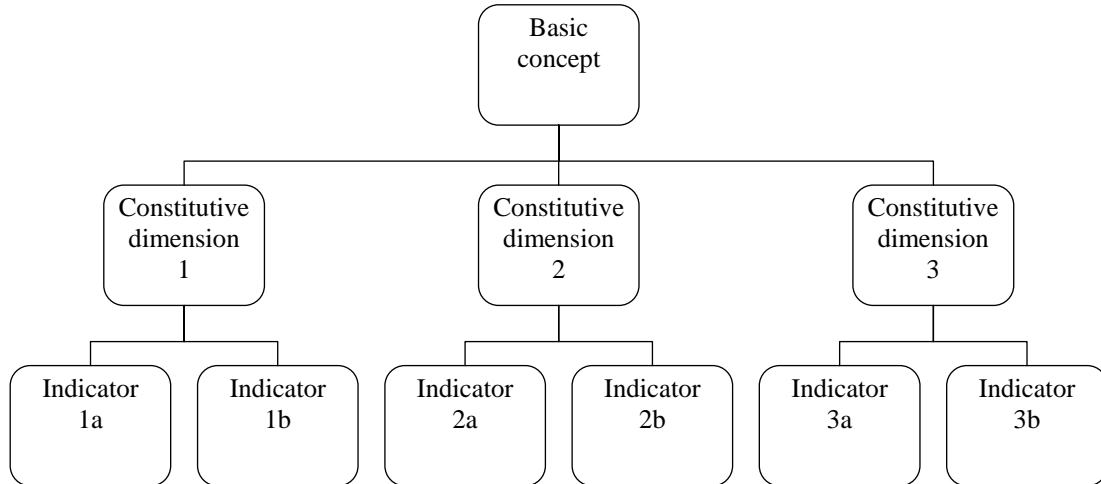
In the present paper, I strove to precise the empirical and conceptual conditions that require judgment (informed and reasoned decision-making) in the collection of comparative political data: unobservable realities, unobserved realities, unexpected realities and conceptual complexities. In general terms, my conclusion is simple: To the extent that we need to rely on judgmental elements in the collection of comparative political data (be it for epistemic, theoretical, or practical reason), we should better recognize the fact, rather than deny it. We should better develop methodological standards that guide the transparent use of judgment, rather than exercise our judgmental faculties in opaque manners that obscure their comparative advantages: the capacity to make sense of vast amounts of disperse and uneven information (informed decision-making) and the public justification of measurement decisions in the light of available evidence (reasoned decision-making).



*Appendix*

---

**FIGURE 1. A SIMPLE CONCEPTUAL HIERARCHY**



**TABLE 1. THE ROLE OF JUDGMENT IN POLITICAL MEASUREMENT**

	ELEMENTS OF MEASUREMENT	METHODOLOGICAL CHALLENGES	ANALYTIC DIMENSIONS	ROLE OF JUDGMENT	BUREAUCRATIC REMEDIES
1	EMPIRICAL PHENOMENA	UNOBSERVABLE REALITIES	<ul style="list-style-type: none"> <li>▪ SYMBOLIC REALITIES (SUCH AS INSTITUTIONS, ACTIONS, WRITTEN AND SPOKEN TEXT)</li> <li>▪ CAUSAL RELATIONS AND COUNTERFACTUALS (AS THEY ARE CONSTITUTIVE TO CONCEPTS)</li> <li>▪ SUBJECTIVE REALITIES (SUCH AS INTENTIONS, PERCEPTIONS, AND MORAL COMMITMENTS)</li> </ul>	<p>COMPREHENSION</p> <p>INFERENCE (ASSUMPTIONS AND ARGUMENTS)</p> <p>COMPREHENSION, RECONSTRUCTION FROM SPEECH AND BEHAVIOR</p>	<p>NONE</p> <p>NONE</p> <p>SELECTION OF OBSERVABLE PROXIES (OR SECONDARY SOURCES).</p>
2	EMPIRICAL INFORMATION	UNOBSERVED REALITIES	<ul style="list-style-type: none"> <li>▪ INCOMPLETE OR UNEVEN INFORMATION</li> <li>▪ INCONSISTENT OR CONTRADICTORY INFORMATION</li> </ul>	<p>FILLING IN GAPS ON THE BASIS OF LOCAL KNOWLEDGE. AD HOC AMENDMENT OF RULES OF IGNORANCE. ADJUDICATION ON THE BASIS OF LOCAL KNOWLEDGE. AD HOC AMENDMENT OF RULES OF ADJUDICATION.</p>	<p>AD HOC AMENDMENT OF RULES OF IGNORANCE</p> <p>AD HOC AMENDMENT OF RULES OF ADJUDICATION.</p>
3	RULE APPLICATION	UNEXPECTED REALITIES	<ul style="list-style-type: none"> <li>▪ UNFORESEEN, HARD CASES, BORDERLINE CASES, THAT UNDERMINE MECHANICAL APPLICATIONS OF CODING RULES.</li> </ul>	<p>APPLICATION OF RULES IN THE LIGHT OF THEIR SPIRIT. AD HOC AMENDMENT OF BOUNDARY RULES.</p>	<p>AD HOC AMENDMENT OF BOUNDARY RULES.</p>
4	CONCEPT STRUCTURE	CONCEPTUAL COMPLEXITY	<ul style="list-style-type: none"> <li>▪ ABSTRACT (RATHER THAN CONCRETE) CONCEPTS: MULTIPLE LEVELS BETWEEN ROOT CONCEPT AND INDICATORS</li> </ul>	<p>SYNTHESIS AND INTEGRATION OF INFORMATION</p>	<p>CONCEPTUAL JUMPING</p>

ELEMENTS OF MEASUREMENT	METHODOLOGICAL CHALLENGES	ANALYTIC DIMENSIONS	ROLE OF JUDGMENT	BUREAUCRATIC REMEDIES
		<ul style="list-style-type: none"> <li>▪ COMPOSITE (RATHER THAN SIMPLE) CONCEPTS: MULTIPLE SUB/DIMENSIONS</li> <li>▪ AGGREGATE (RATHER THAN SINGULAR) REFERENTS: SPATIAL, TEMPORAL, OR SOCIAL AGGREGATION</li> </ul>	<p>SYNTHESIS AND INTEGRATION OF INFORMATION</p> <p>SYNTHESIS AND INTEGRATION OF INFORMATION</p>	<p>LITMUS TESTS</p> <p>RULES OF REPRESENTATION</p>

## References

---

- Bollen, Kenneth A. and Pamela Paxton (2000), "Subjective Measures of Liberal Democracy", *Comparative Political Studies* 33/1 (February): 58-86.
- Brady, Henry E. (2004), "Doing Good and Doing Better: How Far Does the Quantitative Template Get Us?" *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, eds. Henry E. Brady and David Collier (Lanham: Rowman and Littlefield), pp. 53-74.
- Collier, Paul and Anke Hoeffler (2004), "Greed and grievance in civil war", *Oxford Economic Papers* 56: 563-595.
- Elklit, Jørgen and Andrew Reynolds (2005), "A Framework for the Systematic Study of Election Quality", *Democratization* 12/2 (April): 147-162.
- Fish, M. Steven (2006), "Creative Constitutions: How Do Parliamentary Powers Shape the Electoral Arena? *Electoral Authoritarianism: The Dynamics of Unfree Competition*, ed. Andreas Schedler (Boulder and London: Lynne Rienner Publishers), pp. 181-197.
- Gervasoni, Carlos (2008), "Conceptualizing and Measuring Subnational Regimes: An Expert Survey Approach", *Political Concepts* 23, IPSA Committee on Concepts and Methods, Working Paper Series ([www.concepts-methods.org](http://www.concepts-methods.org)).
- Goertz, Gary (2006), *Social Science Concepts: A User's Guide* (Princeton: Princeton University Press).
- Habermas, Jürgen (1981), *Theorie des kommunikativen Handelns: Handlungsrationalität und gesellschaftliche Rationalisierung* (Frankfurt am Main: Suhrkamp).
- Heath, Anthony, Stephen Fisher and Shawna Smith (2005), "The Globalization of Public Opinion Research", *Annual Review of Political Science* 8: 297-333.
- Johnson, James (2006), "Consequences of Positivism: A Pragmatist Assessment", *Comparative Political Studies* 39/2 (March): 224-252.
- King, Gary, Robert O. Keohane and Sidney Verba (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, NJ: Princeton University Press).
- Kurtz, Marcus and Andrew Schrank (2008), "Promises and Perils of Cross-National Datasets: Perceptions, Objective Indicators and 'the Rule of Law'", 104<sup>th</sup> Annual Meeting, American Political Science Association (APSA), Boston, Massachusetts, 28-31 August.
- Landman, Todd and Julia Häusermann (2003), "Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance", University of Essex, Human Rights Centre and Statistical Office of the Commission of the European Union (Eurostat), unpublished report.
- Lieberman, Evan S. (2010), "Bridging in Qualitative-Quantitative Divide: Best Practices in the Development of Historically-Oriented Replication Datasets", *Annual Review of Political Science* 13: 35-59.
- Michell, Joel (2005), "Measurement Theory", *Encyclopedia of Social Measurement*, Volume 2, ed. Kimberly Kempf-Leonard (Oxford: Elsevier Academic Press), pp. 677-682.

- Munck, Gerardo L. and Jay Verkuilen (2002), "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices", *Comparative Political Studies* 35/1 (February): 5-34.
- Neuendorf, Kimberly A. (2002), *The Content Analysis Guidebook* (Thousand Oaks: Sage).
- Pepinsky, Thomas B. (2007), "How to Code", *Political Concepts* 18, IPSA Committee on Concepts and Methods, Working Paper Series ([www.concepts-methods.org](http://www.concepts-methods.org)).
- Piano, Ali, Arch Puddington and Mark Y. Rosenberg (eds.) (2006), *Freedom in the World 2006: The Annual Survey of Political Rights and Civil Liberties* (New York: Freedom House and Lanham: Rowman & Littlefield).
- Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub and Fernando Limongi (2000), *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990* (Cambridge, UK: Cambridge University Press).
- Sala, Brian R., John T. Scott and James F. Spriggs II (2007), "The Cold War on Ice: Constructivism and the Politics of Olympic Figure Skating Judging", *Perspectives on Politics* 5/1 (March): 17-29.
- Sartori, Giovanni (2009a), "Concept misformation in comparative politics", *Concepts and Method in Social Science*, eds. David Collier and John Gerring (New York and London: Routledge, orig. 1970), pp. 13-43.
- Sartori, Giovanni (2009b), "Guidelines for concept analysis", *Concepts and Method in Social Science*, eds. David Collier and John Gerring (New York and London: Routledge, orig. 1984), pp. 97-150.
- Schaffer, Frederic Charles and Andreas Schedler (2007), "What Is Vote Buying?" *Elections for Sale: The Causes and Consequences of Vote Buying*, ed. Frederic Charles Schaffer (Boulder and London: Lynne Rienner Publishers), pp. 17-30.
- Schedler, Andreas (2011), "Concept Formation", *International Encyclopedia of Political Science*, eds. Bertrand Badie, Dirk Berg-Schlosser and Leonardo Morlino (London: Sage Publishers), forthcoming.
- Schedler, Andreas (2012), "The Measurer's Dilemma: Coordination Failures in Cross-National Political Data Collection", *Comparative Political Studies* 45/2 (February): forthcoming.
- Wilson, Andrew (2005), *Virtual Politics: Faking Democracy in the Post-Soviet World* (New Haven: Yale University Press).



## Novedades

---

### DIVISIÓN DE ADMINISTRACIÓN PÚBLICA

- Judith Mariscal y Federico Kuhlmann, *Effective Regulation in Latin American Countries. The cases of Chile, Mexico and Peru*, DTAP-236
- Ma. Amparo Casar, *La otra reforma*, DTAP-237
- Laura Sour y Fredy Girón, *Electoral Competition and the Flypaper Effect in Mexican Local Governments*, DTAP-238
- Laura Sour, *Gender Equity, Enforcement Spending and Tax Compliance in Mexico*, DTAP-239
- Lizbeth Herrera y José Ramón Gil García, *Implementación del e-gobierno en México*, DTAP-240
- Ma. Amparo Casar, Ignacio Marván y Khemvirg Puente, *La rendición de cuentas y el poder legislativo*, DTAP-241
- Sergio Cárdenas, Ignacio Lozano, Miguel Torres y Katsumi Yamaguchi, *Identificando beneficiarios de programas gubernamentales*, DTAP-242
- Sergio Cárdenas, *Obstáculos para la calidad y la equidad: La corrupción en los sistemas educativos*, DTAP-243
- Sergio Cárdenas, *Separados y desiguales: Las escuelas de doble turno en México*, DTAP-244
- María del Carmen Pardo, *Los mecanismos de rendición de cuentas en el ámbito ejecutivo de gobierno*, DTAP-245

### DIVISIÓN DE ECONOMÍA

- Antonio Jiménez, *Notes on the Constrained Suboptimality Result by J. D. Geanakoplos and H. M. Polemarchakis (1986)*, DTE-466
- David Mayer, *Long-Term Fundamentals of the 2008 Economic Crisis*, DTE-467
- Luciana Moscoso, *Labels for Misbehavior in a Population With Short-Run Players*, DTE-468
- Daniel Ángeles y Rodolfo Cermeño, *Desempeño de estimadores alternativos en modelos GARCH bivariados con muestras finitas*, DTE-469
- Antonio Jiménez, *Strategic Information Acquisition in Networked Groups with "Informational Spillovers"*, DTE-470
- Rodolfo Cermeño y Mahetabel Solís, *Impacto de noticias macroeconómicas en el mercado accionario mexicano*, DTE-471
- Víctor Carreón, Juan Rosellón y Eric Zenón, *The Hydrocarbon Sector in Mexico: From the Abundance to the Uncertain Future*, DTE-472
- John Scott, *The Incidence of Agricultural Subsidies in Mexico*, DTE-473
- Alfredo Cuecuecha y John Scott, *The Effect of Agricultural Subsidies on Migration and Agricultural Employment*, DTE-474
- Alejandro Villagómez y Luis Navarro, *Política fiscal contracíclica en México durante la crisis reciente: Un análisis preliminar*, DTE-475

## DIVISIÓN DE ESTUDIOS INTERNACIONALES

- Alejandro Anaya, *Altos niveles de presión transnacional sobre México por violaciones de derechos humanos*, DTEI-190
- Andrea Barrios, *Food Security and WTO Obligations in the Light of the Present Food Crisis*, DTEI-191
- Covadonga Meseguer y Abel Escribà Folch, *Learning, Political Regimes and the Liberalization of Trade*, DTEI-192
- Jorge Chabat, *El narcotráfico en las relaciones México-Estados Unidos: Las fuentes del conflicto*, DTEI-193
- Farid Kahhat y Carlos E. Pérez, *El Perú, Las Américas y el Mundo. Política exterior y opinión pública en el Perú 2008*, DTEI-194
- Jorge Chabat, *La Iniciativa Mérida y la relación México-Estados Unidos: En busca de la confianza perdida*, DTEI-195
- Jorge Chabat, *La respuesta del gobierno de Calderón al desafío del narcotráfico: Entre lo malo y lo peor*, DTEI-196
- Jorge Schiavon y Rafael Velázquez, *La creciente incidencia de la opinión pública en la política exterior de México: Teoría y realidad*, DTEI-197
- Rafael Velázquez y Karen Marín, *Política exterior y diplomacia parlamentaria: El caso de los puntos de acuerdo durante la LX Legislatura*, DTEI-198
- Alejandro Anaya, *Internalización de las normas internacionales de derechos humanos en México*, DTEI-199

## DIVISIÓN DE ESTUDIOS JURÍDICOS

- María Mercedes Albornoz, *Choice of Law in International Contracts in Latin American Legal Systems*, DTEJ-36
- Gustavo Fondevila, *Contacto y control del sistema de informantes policiales en México*, DTEJ-37
- Ana Elena Fierro y Adriana García, *¿Cómo sancionar a un servidor público del Distrito Federal y no morir en el intento?*, DTEJ-38
- Ana Elena Fierro, *Transparencia: Herramienta de la justicia*, DTEJ-39
- Marcelo Bergman, *Procuración de justicia en las entidades federativas. La eficacia del gasto fiscal de las Procuradurías Estatales*, DTEJ-40
- José Antonio Caballero, *La estructura de la rendición de cuentas en México: Los poderes judiciales*, DTEJ-41
- Ana Laura Magaloni, *El ministerio público desde adentro: Rutinas y métodos de trabajo en las agencias del MP*, DTEJ-42
- Carlos Elizondo y Ana Laura Magaloni, *La forma es fondo. Cómo se nombran y cómo deciden los ministros de la Suprema Corte de Justicia de la Nación*, DTEJ-43
- María Mercedes Albornoz, *Utilidad y problemas actuales del crédito documentario*, DTEJ-44
- Gustavo Fondevila, *"Madrinas" en el cine. Informantes y parapolicías en México*, DTEJ-45



## DIVISIÓN DE ESTUDIOS POLÍTICOS

- Francisco Javier Aparicio y Sandra Jessica Ley, *Electoral Institutions and Democratic Consolidation in the Mexican States, 1990-2004*, DTEP-208
- Joy Langston, *Las reformas electorales de 2007*, DTEP-209
- Carlos Elizondo, *La industria del amparo fiscal*, DTEP-210
- María de la Luz Inclán, *Threats and Partial Concessions in the Exhaustion of the Zapatista Wave of Protest, 1994-2003*, DTEP-211
- Andreas Schedler, *Inconsistencias contaminantes. Gobernación electoral y conflicto postelectoral en las elecciones presidenciales de 2006*, DTEP-212
- Andreas Schedler, *Academic Market Failure. Data Availability and Quality in Comparative Politics*, DTEP-213
- Allyson Benton, *Politics and Sector-Specific Stock Market Performance*, DTEP-214
- Andreas Schedler, *The New Institutionalism in the Study of Authoritarian Regimes*, DTEP-215
- Julio Ríos Figueroa, *Institutions for Constitutional Justice in Latin America*, DTEP-216
- Francisco Javier Aparicio y Joy Langston, *Committee Leadership Selection without Seniority: The Mexican Case*, DTEP-217

## DIVISIÓN DE HISTORIA

- Mauricio Tenorio, *Around 1919 and in Mexico City*, DTH-56
- Michael Sauter, *Between Outer Space and Human Space: Knowing Space as the Origin of Anthropology*, DTH-57
- Luis Medina, *Federalismo mexicano para principiantes*, DTH-58
- Mónica Judith Sánchez, *Liberal Multiculturalism and the Problems of Difference in the Canadian Experience*, DTH-59
- Luis Medina, *El Plan de Monterrey de 1855: un pronunciamiento regionalista en México*, DTH-60
- Luis Medina, *La organización de la Guardia Nacional en Nuevo León*, DTH-61
- Luis Medina, *La Comanchería*, DTH-62
- Jean Meyer, *Historia y ficción, hechos y quimeras*, DTH-63
- Ugo Pipitone, *Kerala, desarrollo y descentralización*, DTH-64
- Ugo Pipitone, *Criminalidad organizada e instituciones. El caso siciliano*, DTH-65

## Ventas

El CIDE es una institución de educación superior especializada particularmente en las disciplinas de Economía, Administración Pública, Estudios Internacionales, Estudios Políticos, Historia y Estudios Jurídicos. El Centro publica, como producto del ejercicio intelectual de sus investigadores, libros, documentos de trabajo, y cuatro revistas especializadas: *Gestión y Política Pública*, *Política y Gobierno*, *Economía Mexicana Nueva Época* e *Istor*.

Para adquirir cualquiera de estas publicaciones, le ofrecemos las siguientes opciones:

VENTAS DIRECTAS:	VENTAS EN LÍNEA:
Tel. Directo: 5081-4003 Tel: 5727-9800 Ext. 6094 y 6091 Fax: 5727 9800 Ext. 6314  Av. Constituyentes 1046, 1er piso, Col. Lomas Altas, Del. Álvaro Obregón, 11950, México, D.F.	Librería virtual: <a href="http://www.e-cide.com">www.e-cide.com</a>  Dudas y comentarios: <a href="mailto:publicaciones@cide.edu">publicaciones@cide.edu</a>

## ¡¡Colecciones completas!!

Adquiere los CDs de las colecciones completas de los documentos de trabajo de todas las divisiones académicas del CIDE: Economía, Administración Pública, Estudios Internacionales, Estudios Políticos, Historia y Estudios Jurídicos.



## ¡Nuevo! ¡¡Arma tu CD!!



Visita nuestra Librería Virtual [www.e-cide.com](http://www.e-cide.com) y selecciona entre 10 y 20 documentos de trabajo. A partir de tu lista te enviaremos un CD con los documentos que elegiste.