Kurtz, Marcus J., and Andrew M. Schrank. "Perception and misperception in governance research: Evidence from Latin America." *Is Good Governance Good for Development?*. Ed. Jomo Kwame Sundaram and Anis Chowdhury. London: Bloomsbury Academic, 2012. 71–96. The United Nations Series on Development. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781780932477.ch-004>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 12:23 UTC.

Copyright © United Nations 2012. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

Chapter 4

Perception and misperception in governance research: Evidence from Latin America

MARCUS J. KURTZ AND ANDREW M. SCHRANK

This chapter identifies and takes issue with a potentially troubling development in the study of comparative politics. While political scientists have traditionally deployed *objective* indicators of 'political organization and administrative capacity' (Reynolds 1983: 976), including tax ratios (Organski and Kugler 1980; Benson and Kugler 1998), tax structures (Kling 1959; Krasner 1985), political participation (Przeworski and Sprague 1971), and the nature and extent of public service provision (Migdal 1988; Putnam 1993), they are beginning to import *subjective* indicators of 'governance' from economics (Sandholtz and Gray 2003; Fish 2005; Gerring and Thacker 2005; Blake and Martin 2006; Cameron, Blararu and Burns 2006). The customary distinction between objective and subjective measurement is, by now, less salient than the various differences *among* the subjective or 'perceptions-based' alternatives (Kaufmann, Kraay and Mastruzzi 2005a). Whose perceptions do they capture? Of which institutions or issue areas? And to what effect?

We hold that the leading perceptions-based indicators are poorly explicated, give the perceptions and interests of businesspeople (and foreign businesspeople in particular) undue influence, are riddled with measurement error and all but impossible to interpret, and should therefore be either reconsidered or replaced by *improved* objective indicators that take public sector inputs as well as outputs into account.

We make the case for reconsideration and/or replacement in four principal sections. The first section offers an abbreviated introduction to the perceptions-based literature. The next section develops a critique of the conceptual and operational underpinnings of the leading perceptions-based indicators. The following section draws an important but frequently overlooked distinction between the quality of governance and the quantity of government. The concluding section discusses objective alternatives.

INTELLECTUAL CONTEXT

Economists first brought perceptions-based indicators of corruption, property rights and the rule of law to bear in cross-national growth research in the mid-1990s. For instance, Paolo Mauro obtained national-level indicators of judicial integrity, corruption and bureaucratic quality from commercial risk rating agencies as early as 1995 (Mauro 1995). Arthur Goldsmith deployed an index of property rights developed by the Heritage Foundation more or less concurrently, and turned to the Corruption Perceptions Index developed by Transparency International for the first time a few years later (Goldsmith 1995, 1999). And Alberto Ades and Rafael DiTella introduced the corruption indicators found in the World Economic Forum's *Global Competitiveness Report* at approximately the same time (Ades and DiTella 1997).

By the end of the decade, however, Daniel Kaufmann and his colleagues at the World Bank Institute had come to believe that 'many of these indicators serve as imperfect proxies for one of a smaller number of more fundamental concepts of governance', including 'the rule of law, government effectiveness, and graft' (Kaufmann, Kraay and Zoido-Lobaton 1999a: 1–2), and had therefore decided to develop a series of 'aggregate governance indicators' designed to capture the fundamental concepts in a more transparent and accurate fashion. They aggregated data from a wide array of sources including – but by no means limited to – the aforementioned foundations, nongovernmental organizations (NGOs) and commercial risk rating agencies; developed an unobserved components model; and distilled the aggregate data into six different indicators: voice and accountability, political instability and violence, government effectiveness, regulatory burden, rule of law and control of corruption.

The so-called Worldwide Governance Indicators (WGIs) have gained rave reviews (Gervasoni 2006) and market share in political science, sociology, law, public policy and, of course, economics (see, for example, Apodaca 2004; Fish 2005; Rigobon and Rodrik 2005; Borrmann, Busse and Neuhaus 2006; Lee 2007) – in no small part due to the creativity and entrepreneurship of the authors. And the WGIs are by no means devoid of merit. They are available for more than 150 countries. They cover a ten-year period. They include standard errors as well as individual country scores. And their authors issue frequent and welcome caveats as to their various limitations (see, for example, Kaufmann, Kraay and Mastruzzi 2005b: 41).

In fact, the WGIs arguably provide the best possible case for the perceptions-based measurement of governance. They are, by all accounts,

among 'the most carefully constructed and widely used' (Arndt and Oman 2006: 49) indicators available. And they are, by now, more than a decade old. If they fail to convince, we argue, then the entire project of perceptions-based measurement demands reconsideration – if not abandonment.

THE LIMITS TO PERCEPTIONS-BASED INDICATORS IN PROSPECT: FROM EXPLICATION TO OPERATIONALIZATION

Do the WGIs convince? We are neither the first nor the only scholars to entertain doubts (Arndt and Oman 2006; Bhagwati 2007; Knack 2007; Thomas 2007). Our doubts begin, however, with the explication of the conceptual underpinnings of the measures themselves (see Kurtz and Schrank 2007a, 2007b). Rudolf Carnap famously defined the process of explication as 'the transformation of an inexact, prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*' (Carnap 1962: 3), and offered a number of apposite examples including the self-conscious transformation of the *explicandum* of 'warmth' into the *explicatum* of 'temperature'. According to Carnap, the description of the *explicandum* is a necessary starting point in the process of measurement and comparison, and is therefore no less central to scientific progress than the eventual interpretation and assessment of the *explicatum*. 'Although the *explicandum* cannot be given in precise terms', he argued, 'it should be made as clear as possible by way of informal explanations and examples' (Carnap 1962: 3). Otherwise, parties to a scientific discussion or dispute would inevitably wind up talking past one another.¹

Our discussion therefore starts with the *explicanda*. What are the prescientific concepts in question? While Kaufmann and his colleagues purport to be measuring institutions, and devote a good deal of time to their operationalization, they offer ad hoc and/or inconsistent definitions of their underlying concepts. Take, for example, their measure of the 'rule of law'. They initially define the rule of law as 'respect of citizens and the state for the rules which govern their interactions' (Kaufmann, Kraay and Zoido-Lobaton 1999b: 2). They subsequently abandon their original definition for 'the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence' (Kaufmann *et al.* 2005a: 4). And they eventually decide to deploy a more encompassing, if less discriminating, definition that includes the 'extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence' (Kaufmann, Kraay and Mastruzzi 2006: 4).

They offer neither Carnap's 'informal explanations and examples' nor, for that matter, a rationale for their definitional amendments – let alone a discussion of their implications. Are we to assume that their rule of law indicator (RL) assesses different concepts in different years? Or that the definitional changes have no effect on the indicator's interpretation? And, if the latter, why do they introduce the changes in the first place?

The implications are by no means trivial. Let us take a concrete example. Italy's rule of law score all but collapsed between 1996, when the Italians (1.05) outperformed the Czechs (0.941), Poles (0.867), South Koreans (0.846) and several other members of the Organization for Economic Co-operation and Development (OECD), and 2006, when it found itself relegated to an ignominious position (0.313) beneath Jordan (0.452), Botswana (0.611), Malaysia (0.564) and Uruguay (0.484) – not to mention a number of monarchies, dictatorships, emirates and tax havens (Kaufmann, Kraay and Mastruzzi 2008, Table C5). In fact, Italy's estimated RL scores have declined from their previous values in seven out of ten updates to the WGI's. Why? Have Italian politics and society really changed so dramatically and so consistently over the course of a single decade? Or did Kaufmann and his colleagues de-emphasize the informal relationships and institutions that loom large in the Italian political economy (Putnam 1993) in favour of formal institutions like the police and courts – that are less operative in much of Italy – when they abandoned their original definition of the rule of law? Our point is neither to whitewash Italy's very real governance dilemmas nor to rule out other explanations for the fall. We simply wish to underscore the fact that conceptual shifts – especially those that are neither explained nor justified by their authors – come with real interpretive costs.²

Our concerns are neither frivolous nor pedantic. After all, the rule of law's empirical referent is anything but obvious. Legal scholars like Thom Ringer have not only condemned the 'conceptual anarchy among development theorists, experts, and donor agencies surrounding the very meaning of the expression', but have gone on to wonder whether the 'problem with measuring the success of rule of law reform initiatives is that the parties assessing them may have something quite different in mind to those implementing them' (Ringer 2007: 182; see also Daniels and Trebilcock 2004). Nor are they alone. Political scientists and sociologists have expressed doubts about the concept as well (Stephenson 2000; Kurtz and Schrank 2007b). And Dani Rodrik has recently wondered whether he is 'the only economist guilty of using the term abundantly without having a good fix on what it really means' or simply 'the first one to confess to it' (Rodrik 2007).

Kaufmann and his colleagues appear to be sanguine by way of comparison. While they take comfort in the fact that the various aspects of governance they purport to measure ‘tend to be quite highly correlated across countries’ (Kaufmann, Kraay and Mastruzzi 2007a: 555), they are in fact assuming what needs to be proven – that is, that their correlated indicators are valid measures of governance in the first place. Otherwise the reported correlations are all but devoid of meaning (see also Knack 2007).³

Are the WGIs valid indicators of governance? We will address the question by focusing primarily upon the most recent iteration of their RL, for ‘measurement validity is specifically concerned with whether operationalization and the scoring of cases adequately reflect the concept *the researcher seeks to measure*’ (Adcock and Collier 2001: 529, italics added), and Kaufmann and his colleagues hold that RL in particular captures the ‘norms of limited government’ (Kaufmann *et al.* 2007a: 555) that are central to the literature on institutions and growth (Kaufmann *et al.* 2007a: 561; see also Kaufmann and Kraay 2002: 192).⁴

Table 4.1 recapitulates the definition of the rule of law that appears in the most recent iteration of the WGIs – that is, the concept that Kaufmann *et al.* purport to measure – and includes summary information on the sources, respondents and questions incorporated into the RL. The questions tend to fall into three principal subgroups: crime, property rights and judicial and security institutions. We evaluate RL’s content validity by reflecting upon each subgroup and asking (i) whether key elements are omitted or (ii) inappropriate elements are included in the indicator’s construction (Adcock and Collier 2001: 538).⁵

Crime

RL includes data on both crime in general and a number of specific offences (e.g. kidnapping) or categories of crime (e.g. organized crime). The underlying data sources include expert assessments of crime prevalence rates and trends, household survey data on victimization, and subjective assessments of the costs that criminals impose on business. And the rationale for their inclusion is more or less obvious: a society characterized by the rule of law is unlikely to be crime-ridden and vice versa.

Our concerns therefore derive not from the incorporation of crime data per se but from the manner in which they are incorporated. After all, the data summarized in Table 4.1 are unlikely to present an unbiased portrait of the level and distribution of crime in society. On the contrary, they are likely to give undue weight to crimes that affect business and elite interests.

Table 4.1:
Components of the WGI rule of law indicator

	<i>Rule of law^a Source</i>	<i>Type</i>	<i>Respondents</i>	<i>Questions</i>
Representative	Global Insight Global Risk Service	Commercial information provider	Staff	Losses and costs of crime Kidnapping of foreigners Enforceability of government contracts Enforceability of private contracts
	Economist Intelligence Unit	Commercial information provider	Correspondents	Violent crime Organized crime Fairness of judicial process Enforceability of contracts Speediness of judicial process Confiscation/ expropriation
	Cerberus Intelligence Gray Area Dynamics	Commercial information provider	Staff	Nationalization/ expropriation
	World Economic Forum Global Competitiveness Survey	Nongovernmental organization (with business representation)	Firms (survey)	Common crime imposes costs on business Organ ized crime imposes costs on business Quality of police The judiciary is independent from political influences of members of government, citizens or firms Legal framework to challenge the legality of government actions is inefficient Intellectual property protection is weak Protection of financial assets is weak Tax evasion

Table 4.1 (cont'd)

	<i>Rule of law^a</i> <i>Source</i>	<i>Type</i>	<i>Respondents</i>	<i>Questions</i>
	Gallup World Poll	Commercial polling firm	Households (survey)	Confidence in the police force Confidence in judicial system Have you been a victim of crime?
	Heritage Foundation Index of Economic Freedom	Nongovernmental organization (conservative)	Staff	Property rights
	Cingranelli & Richards Human Rights Database	Academics	Expert codings of US Department of State and Amnesty International reports	Judicial independence
	Institutional Profiles Database	French Ministry of the Economy, Finance, and Industry and Agence Francais de Developpement	Expert assessments	Respect for law in relations between citizens and the administration Security of persons and goods Organized criminal activity (drug trafficking, arms trafficking etc.) Importance of the informal economy Importance of tax evasion in the formal sector Importance of customs evasion (smuggling, underdeclaration etc.) Running of the justice system Security of traditional property rights Security of property rights: formal property rights Security of contracts between private agents

Table 4.1 (cont'd)

	<i>Rule of law^a Source</i>	<i>Type</i>	<i>Respondents</i>	<i>Questions</i>
				Government respect for contracts Settlement of economic disputes: justice in commercial matters Intellectual property Arrangements for the protection of intellectual property Agricultural sector: Security of rights and property transactions.
	Political Risk Service International Country Risk Guide	Commercial information provider	Staff	Law and order: The law subcomponent is an assessment of the strength and impartiality of the legal system, while the order subcomponent is an assessment of popular observance of the law (assessed separately)
	Business Environment Risk Intelligence Financial Ethics Index	Commercial information provider	Panel of experts	Financial fraud Money laundering Organized crime
	US Department of State Trafficking in Persons Report	USDOS	Expert assessments	Human trafficking
	Global Insight Business Conditions and Risk Guide	Commercial information provider	Staff	Judicial independence: An assessment of how far the state and other outside actors can influence and distort the legal system. This will determine the level of legal impartiality investors can expect Crime: How much of a threat businesses face from crime such as kidnapping, extortion, street violence and burglary

Table 4.1 (cont'd)

	<i>Rule of law^a Source</i>	<i>Type</i>	<i>Respondents</i>	<i>Questions</i>
	African Development Bank Country Policy and Institutional Assessments	African Development Bank	Country economists	Property rights
Non-representative	Afrobarometer	University/NGO collaboration	Household survey	<p>Based on your experiences, how easy or difficult is it to obtain help from the police when you need it?</p> <p>Over the past year, how often, if ever, have you or anyone in your family feared crime in your own home?</p> <p>Over the past year, how often, if ever, have you or anyone in your family had something stolen from your house?</p> <p>Over the past year, how often, if ever, have you or anyone in your family been physically attacked?</p> <p>Trust in courts</p>
	Asian Development Bank Country Policy and Institutional Assessments	Asian Development Bank	Country economists	Rule of law
	Business Environment and Enterprise Performance Surveys	World Bank & European Bank for Reconstruction and Development	Survey (firms)	<p>Fairness, honesty, enforceability and quickness of the court system</p> <p>How problematic is crime for the growth of your business</p> <p>How problematic is judiciary for the growth of your business</p>
	Bertelsmann Transformation Index	Nongovernmental organization	Staff	Rule of law Private property

Table 4.1 (cont'd)

	<i>Rule of law^a</i> <i>Source</i>	<i>Type</i>	<i>Respondents</i>	<i>Questions</i>
	Freedom House Countries at the Crossroads	Nongovernmental organization	Staff and consultants	Rule of law
	World Bank Country Policy and Institutional Assessments	World Bank	Country economists	Property rights
	Freedom House	Nongovernmental organization	Staff and consultants	Judicial framework and independence
	Global Integrity Index	Nongovernmental organization	Local country experts and peer review	Executive accountability Judicial accountability Rule of law Law enforcement
	International Fund for Agricultural Development (IFAD) Rural Sector Performance Assessments Latinobarometer	IFAD Nongovernmental organization	Country economists Survey (household)	Access to land Access to agricultural water Trust in police and judiciary Crime victimization
	Institute for Management Development's World Competitiveness Yearbook	Education organization	Survey (business)	Tax evasion is a common practice in your country Justice is not fairly administered in society Personal security and private property are not adequately protected Parallel economy impairs economic development in your country Patent and copyright protection is not adequately enforced in your country
	Americas Barometer	Vanderbilt University	Survey (household)	Crime victimization Trust in police, courts and judiciary

a Measuring the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police and the courts, as well as the likelihood of crime and violence.

Source: Adapted from Kaufmann, Kraay and Mastruzzi 2009, especially Table B5. Representative sources are available for most countries and therefore weigh more heavily in the final indicators.

While the surveys undertaken by the World Bank and the World Economic Forum (WEF) are self-consciously designed to tap the costs crime imposes on business (see, for example, Kaufmann *et al.* 2008, Table B5), the household surveys undertaken by Gallup and the various regional barometers are likely to prioritize the interests of the elite for a less obvious reason. Respondents in crime victimization surveys are known to overreport property crimes that tend to affect the better off and underreport violent crimes that tend to affect their less fortunate compatriots (Bergman 2006: 221).

Thus, the crime data incorporated into RL are best portrayed neither as proxies for the rule of law writ large nor as random variables but as indicators of the elite's ability to insulate itself from what Alejandro Portes and Brian Roberts (2005: 67) have labelled the 'forced entrepreneurship' of the poor in the Latin American context. According to Portes and Roberts, the relentless onslaught of debt, austerity and free-market reform has provoked the unprecedented growth not only of poverty and inequality but also of anger and resentment in Latin American cities. 'Property crime may rise in these contexts,' they maintain, 'as some members of the subordinate classes take matters into their own hands in order to redress both absolute and relative deprivation.'⁶

Our point is most assuredly neither to rationalize property crimes nor to deny their relationship to the rule of law but to underscore the fact that they are neither inclusive of the full array of crime in most societies nor random occurrences, and that RL therefore *omits* or *obscures* important elements of the rule of law (i.e. violent crime, sex crime, violations of labour and environmental law etc.) and in all likelihood *includes* – albeit indirectly – the various correlates of crime at the country level (e.g. poverty, inequality, the age distribution of the population, collective efficacy etc. (see, for example, Sampson *et al.* 1997)).

Property rights

RL includes data on property rights and threats to their sanctity (i.e. expropriation) as well as information on a number of specific categories of property rights (i.e. intellectual property). The property rights indicators are drawn from commercial risk rating agencies, firm-level surveys and inter- and nongovernmental organizations. But the rationale for their inclusion is unclear, for Kaufmann *et al.*'s definition of the rule of law neither made reference to property rights until 2008 (cf. Kaufmann *et al.* 2007b: 4, 2008: 7) nor militates against their compromise or transgression by *legal* means in any event.

Take, for example, eminent domain provisions in US property law. They give the government the authority to override legitimate property claims in a host of well-defined situations, and are therefore perfectly consistent with the rule of law, but nonetheless give rise to virulent opposition in practice, and are therefore likely to undercut the RL score if used too frequently.

Our objection is readily addressed, however, neither by asserting the infrequency of eminent domain nor by retroactively including a reference to property rights in the definition of RL, for the expropriation and confiscation of property are frequently deployed in the war *against* corruption and crime, and their use is therefore no less consistent with the growth than the absence of the rule of law. Park Chung Hee's campaign against the 'illicit accumulators' in South Korea offers a case in point (Schrank 2007), and more recent examples are readily available in Africa, Asia and Latin America (see, for example, *The Economist* 1981; Arnold 1999).

But one need not travel to the developing world to witness the confiscation and expropriation of assets deployed in defence, rather than contravention, of the rule of law. The Racketeer Influenced and Corrupt Organizations (RICO) Act offers a no less apposite North American example. After all, RICO gives US government prosecutors the right to seize the assets of allegedly corrupt or criminal organizations before they have attained a conviction – and thereby threatens the economic lifeblood of suspect enterprises. A number of observers have given RICO credit for the demise of the Mafia in the 1990s (Laurence 2004). And James Jacobs and Lauryn Gouldin (1999: 169) have therefore labelled the act 'the most important substantive anti-organized crime statute in history'.

The statute has not, however, been free of criticism. While law enforcement officials are justifiably fond of RICO, and favourably disposed towards asset forfeiture in particular, their critics decry 'policing for profit' (Blumenson and Nilson 1998) and worry that the law not only gives the government a 'license to steal' (Levy 1996) from legitimate businesses but also constitutes a 'nuclear deterrent to rational negotiations' (Arkin in *Newsweek* 1989). Is the cure of RICO worse than the disease of corruption and crime? The answer is anything but obvious, for the RICO experience suggests not only that the same institution can simultaneously militate in favour and against the rule of law but also that property claims themselves are, by their very nature, subjective and controversial – even in the advanced industrial countries.⁷

They are arguably more controversial in the developing world, however, and are therefore decidedly ill-equipped to play a meaningful role in perceptions-based indicators of the rule of law. 'A government that evicts squatters will in all likelihood be portrayed as a threat to private property

by the squatters and a bulwark against expropriation by the landlords', we have argued, 'and the problem is likely to be compounded by the fact that in much of the developing world this year's squatters are likely to be next year's landlords and vice versa' (Kurtz and Schrank 2007b).⁸

Nor are the examples of allegedly ill-gotten gains in the United States and squatters in the developing world unique. The problem of ambiguous property claims is widespread and will, in all likelihood, be aggravated by the growing salience of intangible or intellectual property in the future (see, for example, Evans 1997), for different states and societies have radically different conceptions of the appropriate length and scope of patent, copyright and trademark protection.

Judicial and security institutions

RL also includes questions on courts, contract enforcement and the police. The 'courts, cops and contracts' cluster, as we call it, incorporates data from firm and household surveys as well as expert assessments; tracks the government's ability and willingness to ensure the safety and security of the population; and plays an almost indisputable part in the assessment of the rule of law.

We nonetheless wonder why Kaufmann and his colleagues simultaneously omit or de-emphasize a series of regulatory agencies that are no less central to the security and well-being of the population, including, but by no means limited to, the tax, labour, food, drug, health and safety inspectorates; banking and financial overseers; and environmental monitors. Regulatory authorities are involved in *law enforcement*, after all, and their virtual exclusion from RL arguably says more about the types of laws valued by the indicator's authors and advocates than about their importance to the concept itself.⁹

In short, RL appears to suffer from a pronounced and systematic pro-business bias. The crimes covered, the institutions included and the interests served all point in the same direction – as do the interests, institutions and crimes overlooked. And the biases will arguably prove not only threatening but also fatal to content validity when aggravated by what, for lack of a better term, we will call sampling error. After all, the bulk of the data incorporated into the measure – approximately two-thirds of the 'representative' sources that are weighted most heavily in the final indicator (see Table 4.1) – are derived from either private investors or their commercial advisers.¹⁰

Kaufmann and his colleagues (2007a: 556) admit that pro-business biases are possible in theory but belittle their impact in practice. We have already

Table 4.2:
Do businesspeople and their neighbours perceive government in the same way?

Question	Responses	Businessperson
Trust in the judiciary	1 = much trust	Odds ratio = 1.15 ($p < .010$)
Trust in the police	2 = some trust	Odds ratio = 1.25 ($p < .001$)
	3 = little trust	
	4 = no trust	

Note: Self-identified businesspeople are coded 1, others are coded 0, country dummies are suppressed and odds ratios for businesspeople are presented next to their parenthesized p values. Non-responses and ‘don’t know’ are dropped.

Source: The data are from Latinobarometer (2005); the more recent data available to Kaufmann, Kraay and Mastruzzi are not publicly available.

established that reported crime victimization varies by social class. Do business perceptions of other aspects of the rule of law differ substantially as well? Fortunately, the Latinobarometer data employed by Kaufmann and his colleagues invite a direct test.

Table 4.2 includes the results of two ordered logistic regression models. The dependent variables are the answers to the Latinobarometer questions about trust in the judiciary and police incorporated into the RL for the year 2005. They are coded from 1 (much trust) to 4 (no trust). They are regressed on an indicator variable that assumes the value of 1 whenever the respondent is a businessperson and a series of country dummies. And they underscore the differences among businesspersons who consistently evince less faith in police and judicial institutions than their compatriots.¹¹

A defender of the WGIs might rebut charges of bias by portraying businesspersons and their advisers as particularly desirable data sources whose opinions should loom large in their indicators. Private investors not only have more experience with institutions like the police and courts than the average citizen, they might argue, but are also in the vanguard of the struggle for growth and development in market societies. But non-businesspersons are neither ignorant nor unimportant to the process of growth and development; on the contrary, they make countless individual decisions – about everything from personal savings and voting to the treatment and education of their children – that aggregate into important societal outcomes. Such decisions not only influence but are influenced by the broader sociopolitical context, however, and a governance indicator worthy of its name will therefore take the ideas and interests of non-businesspersons seriously.

In short, we have argued not only that the concept of rule of law is poorly explicated but that the RL itself is invalidated by skewed questions addressed towards a biased sample of respondents. The consequences are particularly acute when RL is incorporated into growth regressions as an exogenous variable, for a positive coefficient could reflect any combination of at least four different underlying processes: first, a positive relationship between *actual rule of law* (e.g. crime control, confidence in the courts) and growth; second, a positive relationship between *some but not all aspects of the rule of law* and growth; third, a positive relationship between the *underpinnings of the rule of law* (e.g. equity and opportunity) and growth; or fourth, *herd behaviour* on the part of investors who receive their advice from the same consultants and risk rating agencies and thereby animate growth regardless of the so-called fundamentals.

LIMITS TO PERCEPTIONS-BASED INDICATORS IN PRACTICE: QUALITY VERSUS QUANTITY

The WGIs are premised upon the liberal premise that ‘he who governs best, governs least’ (Kurtz and Schrank 2007a: 543). After all, Kaufmann and his colleagues explicitly reject the Weberian ideal of ‘legal-rational’ authority for a definition of administrative efficacy rooted in ‘the norms of limited government that protect private property from predation by the state’ (Kaufmann *et al.* 2007a: 555).¹² And consumers of their data tend to conclude that states in developing countries need to be ‘not only smaller, but smarter’ (Lora and Panizza 2003: 135).

In fact, the WGIs and their consumers tend to assume that there is an inverse relationship between the quality and the quantity of government. But the WGIs themselves make no effort to distinguish the quality of governance from the quantity (or size) of government, let alone to assign relative weights to each. Governance scores presumably rise with decreases in crime, corruption and their correlates, but whether those decreases are, in fact, the products of better (and presumably less) governance or more governance (whether ‘better’ or not) is by no means clear. Sometimes, governments really can make a difference by throwing more resources (e.g. police, teachers, judges) at a problem; often, they have no alternative.

Figure 4.1 is illustrative. It plots an admittedly crude indicator of national well-being – that is, the Human Development Index (HDI) constructed by the United Nations Development Programme (UNDP 2007) – against a similarly crude indicator of the size of government – that is, the share of the

Figure 4.1:
Human development by public employment (2005)

Source: UNDP (2007), Table 1; Carrizosa (2007), Table 1.

labour force in the public sector (Carrizosa 2007) – for 18 non-communist Latin American countries in 2005.

We draw three conclusions from the scatter plot. First, it is not clear that smaller governments are better governments – at least not if ‘better’ is associated with well-being. If anything, the opposite seems to be the case, for the bivariate correlation between the HDI and the size of the public sector is 0.68. Second, while there does seem to be a ‘small government’ path to human development, as exemplified by Chile, it is decidedly uncommon. Seven Latin American countries have HDIs of 0.8 or above. Six of those countries have the largest public sectors in the region. And, third, a large public sector has all the hallmarks of a ‘sufficient condition’ (Ragin 2000: Chapter 9) for high human development. While Chile is found in the north-western quadrant of the scatter plot, and thus suggests that a large public sector is not a necessary prerequisite for a high HDI, there are no countries in the south-eastern quadrant, suggesting that a Latin American country with a large public sector will invariably guarantee its citizens a reasonable degree of human development.

Of course, Kaufmann and his colleagues might respond to the first conclusion by questioning whether the countries in the north-eastern quadrant are, in fact, well governed. If their public sectors are approximately twice as large as the Chilean public sector, yet their human development performance is no better, after all, there is reason to believe that they are

Figure 4.2:
HDI by 'rule of law'

Source: UNDP (2007); Kaufmann *et al.* (2008).

getting less ‘bang’ for the proverbial ‘buck’ – and that their quality lags accordingly. But from a methodological and policy perspective, this is a bug and not a feature of the Kaufmann *et al.* approach, for both the difference between quality and quantity and the corresponding policy implications are obscured by their measure. Consider Figure 4.2, which plots the HDI against the WGI RL score for 2005.

One finds a similarly high (if decidedly less neat) correlation but a much murkier picture. Once again, Chile is in the north-eastern quadrant but it is suddenly joined by Uruguay and Costa Rica. Meanwhile, a number of high human developers have lower governance scores than El Salvador. How do we make sense of this picture? Ironically, the best way of doing so is returning to the objective data in Figure 4.1, where we can clearly discern two different paths to human development: the small government Chilean road and the big government road taken by the other regional success stories. While in Figure 4.2 Costa Rica and Uruguay appear to have much in common with Chile, Figure 4.1 reveals that they are actually much more comparable to Argentina and Mexico.

In fact, Figure 4.1 suggests that a large public sector is a sufficient cause of high human development and leads to an obvious hypothesis. The Chilean road entails limited extraction and efficient distribution; the big government alternative presupposes more aggressive extraction and less

efficient distribution; and the absence of cases in the south-eastern quadrant derives from the fact that any government willing and able to extract at that level is obliged to effectuate some measure of distribution (however inefficiently) as well.¹³

What, if anything, are the policy implications? While Kaufmann *et al.* would presumably hold out the Chilean path as the rational one to follow, they are largely blind to structural constraints. Efforts to extract and efforts to distribute require distinct social structures and skill sets and are thus differentially possible in different types of societies. And the Chilean road may well demand more (or different) state capacity than most developing countries are able to muster. Does this mean they should let their people suffer in silence (or worse)? Perhaps not, for the evidence in Figure 4.1 suggests that aggressive extraction with less efficient distribution provides a more common – if by no means easily attainable – path to improved human development and may not only be more accessible to public officials who are unwilling or unable to pursue the Chilean path but also more acceptable to their citizens than the status quo.

CONCLUSION: BRINGING OBJECTIVITY BACK IN?

Where do we go from here? We begin to provide an answer by briefly reviewing the experience of one of the more successful subjective measures in the history of the social and medical sciences: self-reported health status. A quarter of a century has passed since Jana Mossey and Evelyn Shapiro first realized that self-reported health status provided a better predictor of seven-year survival rates among the Canadian elderly than either data drawn from medical records or self-reports of specific conditions (Mossey and Shapiro 1982), and much has been learned in that time. Subjective assessments of global health status are by now known to (i) be highly correlated with other indicators (or correlates) of health and well-being, (ii) add enormous explanatory power to multivariate models of mortality (Idler and Benyamini 1997) and (iii) suffer from conceptual limitations that sharply circumscribe their practical utility (Krause and Jay 1994).

Medical professionals know that individuals who label their own health ‘poor’ are likely to die sooner than people who consider themselves fit (Idler and Benyamini 1997), but they do not know why they do so and are therefore unable to act upon their knowledge. Some suspect that self-reports are more inclusive than objective data, tap undiagnosed diseases or capture the effects of co-morbidity. Others think they capture trajectories

rather than levels. Some point to underlying correlations with family history, socio-economic status or behavioural characteristics, and still others to self-fulfilling prophecies. But unless and until an actual causal mechanism is adduced, medical professionals will continue to act not upon subjective but upon objective data when choosing treatment.

Perceptions-based measures of governance have much in common with subjective assessments of health status. They are, after all, highly correlated with each other and with the presumed covariates of good governance (e.g. GDP per capita, school attainment). They, too, add explanatory power to multivariate models. And they also suffer conceptual shortcomings that undercut their practical utility. Take, for example, the rule of law. A positive RL coefficient in a cross-national growth regression is difficult to interpret, let alone act upon. It could reflect the influence of one or more of the actual inputs to RL (e.g. crime, courts, contracts), which are themselves endogenous and are therefore not readily altered by public policy. Or it could betray the impact of the direct effects of the social and economic conditions that influence those inputs (i.e. per capita income, social capital etc.), which are arguably even less susceptible to policy manipulation. It could be a product of pure measurement error of the sort admitted (but perhaps underestimated) by the indicator's authors. Or it could be the product of herd behaviour among investors who purchase their information from the indicator's underlying sources. Unless and until we have clear answers to these questions, however, the WGIs will be of limited practical utility.

The problem is aggravated, we believe, by the implicitly contradictory targets embedded in the indicators and sub-indicators that go into the WGIs. After all, the sources employed by Kaufmann and his colleagues reward governments for policies that are almost certainly in tension with each other. What, then, should developing country policymakers who hope to improve their WGI scores *do*? Outlaw the expropriation of assets in order to maximize RL or adopt laws like RICO and use them to expropriate corrupt businesspeople and politicians so as to gain better scores on the control of corruption (CC) indicator? Extract and invest resources in schools and roads so as to bolster 'government effectiveness' (GE) or cut taxes in order to ensure 'regulatory quality' (RQ)? Clamp down on protest so as to ensure 'political stability' (PS) or give the protesters and the press a free pass in an effort to maximize 'voice and accountability' (VA)? Almost every potential solution aggravates another problem, and the WGIs therefore punish poor countries for their very poverty. After all, if poor countries could solve their social and economic problems, they would not be poor – let alone worried about their governance scores – in the first place.

What, then, is to be done? The medical professionals who collect and analyse self-reported health data ultimately turn to objective alternatives when diagnoses and treatments are necessary. Unfortunately, however, objective governance data are not readily available. The traditional proxies – for example, tax ratios and the like – have been largely discredited for one reason or another. And nobody has seen fit to invest in the creation of alternatives. We think such an effort is long past due, however, and would yield important pay-offs not only for social scientists and policymakers who are sceptical of perceptions-based measures but also, ironically, for their defenders, for the existing subjective indicators arguably cry out for the legitimating aura of an objective benchmark.

What would objective governance indicators look like? First, they would, of necessity, be issue-specific. While subjective health data are ‘global’ in nature, the objective data that guide diagnosis and treatment are necessarily circumscribed to particular health problems. Similar specification would be required for governance data, especially in light of the fact that governance is known to vary across issue areas and sectors within countries as well as among countries more generally (see, for example, Johnson 1982). Second, they would focus on issues with measurable outputs: education, health care provision and postal delivery come immediately to mind (see Putnam 1993). Third, they would incorporate data on *inputs* as well as outputs. One problematic feature of the WGIs is their almost complete inattention to the problem of opportunity cost. A government that bankrupted itself in a successful war against corruption would presumably see its corruption control score improve – at least in the short run. But it is not clear that it would have made a wise choice in doing so. And, finally, it would differentiate the efficacy of governance – that is, the enforcement of the rules of the game – from the quality of policymaking – that is, the creation of the rules of the game – as well as from the sociocultural context that underpins both the rules and their enforcement. It would thus demand independent data on inputs and outputs as well as policies and contextual factors that might affect both.

We can do no more than provide a brutally abbreviated example at present. Neither available data nor the space we have left would tolerate more. But we will try to illustrate the *sort* of approach we think desirable by examining the issue of health care provision in Latin America – which, at least in theory, meets the aforementioned criteria. It is not only narrow in scope, at least when compared to ‘global’ issues like the rule of law, but is widely regarded as a key product of state capacity (Caldwell 1986). In addition, measures of inputs as well as outputs are available for this

Figure 4.3:
Life expectancy by public health spending per capita c. 2005

Source: World Bank (2008); PAHO 2007 (Table 7).

indicator. We treat life expectancy at birth as an indicator of health output. We treat the level of public health care spending per capita as an indicator of the quantity of health inputs. And we assume that at least part of the unexplained variation in output is due to quality differences.

The data in Figure 4.3 suggest several interesting conclusions about governance. First, they fly directly in the face of the widespread notion that ‘limited government’, in Kaufmann *et al.*’s formulation, is necessarily the best government. While the cross-sectional data permit no causal claims, they are at least consistent with the idea that public health care spending promotes, rather than undercuts, life expectancy in Latin America. In fact, the absence of countries in the north-western quadrant raises the possibility that a high level of public health care spending is a *necessary* condition of very high levels of life expectancy. Second, they suggest that quality matters too. Spending alone accounts for only half of the variation in health care output. At least part of the residual may be a product of quality differentials. And finally, they suggest that objective governance indicators could, with a good deal of additional effort, provide a valuable complement or even alternative to the existing perceptions-based measures. One would not only need longitudinal as well as cross-sectional data from a larger number of countries but would also have to deal with questions of endogeneity and selection. But these issues arise with subjective data as well – and there they

are compounded by issues of conceptual ambiguity, systematic and random measurement error, and causal indeterminacy.

A thoroughgoing collective effort to aggregate and disseminate objective data would not only help researchers combat such problems but would also provide an alternative if such a campaign were to fail. Data on inputs and outputs are available across a wide array of issue areas in a large number of developing countries. It is high time that intergovernmental organizations began to aggregate them and make them available to policymakers, scholars and activists around the world.

REFERENCES

- Adcock, Robert, and David Collier (2001). Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review*, 95 (3): 529–46.
- Ades, Alberto, and Rafael DiTella (1997). National Champions and Corruption: Some Unpleasant Interventionist Arithmetic. *Economic Journal*, 107 (443): 1023–42.
- Apodaca, Claire (2004). The Rule of Law and Human Rights. *Judicature*, 87 (6): 292–99.
- Arndt, Christiane, and Charles Oman (2006). *Uses and Abuses of Governance Indicators*. Paris: OECD.
- Arnold, Wayne (1999). Tough Guy at the Loan Office; Backed by I.M.F., a Reformer Tackles Indonesia's Barons. *The New York Times*, July 28.
- Benson, Michelle, and Jacek Kugler (1998). Power Parity, Democracy, and the Severity of Internal Violence. *Journal of Conflict Resolution*, 42 (2): 196–209.
- Bergman, Marcelo (2006). Crime and Citizen Security in Latin America: The Challenges for New Scholarship. *Latin American Research Review*, 41 (2): 213–27.
- Bhagwati, Jagdish (2007). Comments, FT Economists' Forum. *Financial Times*, November 7. Available at <http://blogs.ft.com/wolfforum/2007/11/why-plutocracy-.html#comments> [accessed 21 February 2012].
- Blake, Charles, and Christopher Martin (2006). The Dynamics of Political Corruption: Re-examining the Influence of Democracy. *Democratization*, 13 (1): 1–14.
- Blumenson, Eric, and Eva Nilson (1998). Policing for Profit: The Drug War's Hidden Economic Agenda. *University of Chicago Law Review*, 65 (1): 35–114.
- Borrmann, Axel, Matthias Busse and Silke Neuhaus (2006). Institutional Quality and the Gains from Trade. *Kyklos*, 59 (3): 345–58.
- Byerly, Henry (1972). Realist Foundations of Measurement. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Chicago: University of Chicago Press, pp. 375–82.
- Caldwell, John (1986). Routes to Low Mortality in Poor Countries. *Population and Development Review*, 12 (2): 171–220.
- Cameron, Maxwell, Ana-Maria Blanaru and Lesley Burns (2006). Constitutional Frameworks and the Rule of Law: Between Delegative Democracy and Elective

- Dictatorship. Paper presented to the American Political Science Association, Philadelphia, September.
- Carnap, Rudolf (1962). *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- Carrizosa, Mauricio (2007). Public Sectors in Latin America: How Big Are They? *En Breve*, 108. July.
- Daniels, Ronald, and Michael Trebilcock (2004). The Political Economy of Rule of Law Reform in Developing Countries. *Michigan Journal of International Law*, 26: 99–108.
- Dorsey, Christine (2001). Nevada Awaits Norton's Attention. *Las Vegas Review-Journal*, August 20.
- Dumont, Richard, and William Wilson (1967). Aspects of Concept Formation, Explication, and Theory Construction in Sociology. *American Sociological Review*, 32 (6): 985–95.
- The Economist* (1981). Chile: Capitalists Gaffed. December 12.
- Evans, Peter (1997). The Eclipse of the State? Reflections on Stateness in an Era of Globalization. *World Politics*, 50 (October): 62–87.
- Fish, M. Steven (2005). The Hazards of Half Measures: Perestroika and the Failure of Post-Soviet Democratization. *Demokratizatsiya*, 13 (2): 241–53.
- Gerring, John, and Strom Thacker (2005). Do Neoliberal Policies Deter Corruption? *International Organization*, 59 (Winter): 233–54.
- Gervasoni, Carlos (2006). Data Set Review: The World Bank's Governance Indicators (1996–2004). *APSA-CP Newsletter*, 17 (1): 17–20.
- Goldsmith, Arthur (1995). Democracy, Property Rights and Economic Growth. *Journal of Development Studies*, 32 (2): 157–74.
- Goldsmith, Arthur (1999). Slapping the Grasping Hand: Correlates of Political Corruption in Emerging Markets. *American Journal of Economics and Sociology*, 58 (4): 865–83.
- Hempel, Carl (1956). Logical Appraisal of Operationism. *Scientific Monthly*, 79: 215–20.
- Idler, Ellen, and Yale Benyamini (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 38 (1): 21–37.
- Jacobs, James, and Lauryn Gouldin (1999). Cosa Nostra: The Final Chapter? *Crime and Justice*, 25: 129–99.
- Johnson, Chalmers (1982). *MITI and the Japanese Miracle: The Rise of Industrial Policy*. Berkeley, CA: University of California Press.
- Kaufmann, Daniel, and Aart Kraay (2002). Growth without Governance. *Economia*, 3 (1): 169–229.
- Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobaton (1999a). *Aggregating Governance Indicators*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobaton (1999b). *Governance Matters*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2005a). *Measuring Governance Using Cross-Country Perceptions Data*. Washington, DC: World Bank Institute.

- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2005b). *Governance Matters IV*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2006). *Governance Matters V*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2007a). Growth and Governance: A Reply. *Journal of Politics*, 69 (2): 555–62.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2007b). *Governance Matters VI*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2007c). *The Worldwide Governance Indicators Project: Answering the Critics*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2007d). Growth and Governance: A Rejoinder. *Journal of Politics*, 69 (3): 570–2.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2008). *Governance Matters VII*. Washington, DC: World Bank Institute.
- Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2009). *Governance Matters VIII*. Washington, DC: World Bank Institute.
- Kling, Merle (1959). Taxes on the ‘External’ Sector: An Index of Political Behavior in Latin America? *Midwest Journal of Political Science*, 3 (2): 127–50.
- Knack, Stephen (2007). Measuring Corruption: A Critique of Indicators in Eastern Europe and Central Asia. *Journal of Public Policy*, 27 (3): 245–91.
- Krasner, Stephen (1985). *Structural Conflict: The Third World against Global Liberalism*. Berkeley, CA: University of California Press.
- Krause, Neal, and Nina Jay (1994). What Do Global Self-Rated Health Items Measure? *Medical Care*, 32 (9): 930–42.
- Kurtz, Marcus, and Andrew Schrank (2007a). Growth and Governance: Models, Measures, and Mechanisms. *Journal of Politics*, 69 (2): 538–54.
- Kurtz, Marcus, and Andrew Schrank (2007b). Growth and Governance: A Defense. *Journal of Politics*, 69 (2): 563–9.
- Laurence, Charles (2004). Mafia’s ‘Last Godfather’ Goes on Trial for Murder and Racketeering. *Sunday Telegraph*, May 16.
- Lee, Cheol-Sung (2007). Labor Unions and Good Governance: A Cross-National, Comparative Analysis. *American Sociological Review*, 72 (4): 585–609.
- Levy, Leonard (1996). *A License to Steal: The Forfeiture of Property*. Chapel Hill, NC: UNC Press.
- Lora, Eduardo, and Ugo Panizza (2003). The Future of Structural Reform. *Journal of Democracy*, 14 (2): 123–37.
- Mauro, Paolo (1995). Corruption and Growth. *Quarterly Journal of Economics*, 110 (3): 681–712.
- Migdal, Joel (1988). *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton, NJ: Princeton University Press.
- Mossey, Jana, and Evelyn Shapiro (1982). Self-Rated Health: A Predictor of Mortality among the Elderly. *American Journal of Public Health*, 72 (8): 800–8.
- Newsweek* (1989). The RICO Weapon: Too Harsh for Wall Street? January 2.
- Organski, A.F.K., and Jacek Kugler (1980). *The War Ledger*. Chicago: University of Chicago Press.

- Pan American Health Organization (PAHO) (2007). *Health in the Americas*. Washington, DC: PAHO.
- Portes, Alejandro, and Brian Roberts (2005). The Free Market City: Latin American Urbanization in the Years of the Neoliberal Experiment. *Studies in Comparative International Development*, 40 (1): 43–82.
- Przeworski, Adam, and Joseph Sprague (1971). Concepts in Search of Explicit Formulation: A Study in Measurement. *American Political Science Review*, 15 (2): 183–218.
- Putnam, Robert (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- Ragin, Charles (2000). *Fuzzy Set Social Science*. Chicago: University of Chicago Press.
- Reynolds, Lloyd (1983). The Spread of Economic Growth to the Third World: 1850–1980. *Journal of Economic Literature*, 21 (3): 941–80.
- Rigobon, Roberto, and Dani Rodrik (2005). Rule of Law, Democracy, Openness and Income: Estimating the Interrelationships. *Economics of Transition*, 13 (3): 533–64.
- Ringer, Thom (2007). Development, Reform, and the Rule of Law: Some Prescriptions for a Common Understanding of the ‘Rule of Law’ and Its Place in Development Theory and Practice. *Yale Human Rights and Development Law Journal*, 10: 78–208.
- Rodrik, Dani (2007). What Does the ‘Rule of Law’ Mean? Weblog entry. November 13. Available at http://rodrik.typepad.com/dani_rodriks_weblog/2007/11/what-does-the-r.html [accessed 27 March 2012].
- Salmon, Wesley (1999). The Spirit of Logical Empiricism: Carl G. Hempel’s Role in Twentieth-Century Philosophy of Science. *Philosophy of Science*, 66 (3): 33–50.
- Sampson, Robert, Stephen Raudensbuh and Felton Earls (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, 277: 918–24.
- Sandholtz, Wayne, and Mark Gray (2003). International Integration and National Corruption. *International Organization*, 57 (Autumn): 761–800.
- Schrank, Andrew (2007). Asian Industrialization in Latin American Perspective: The Limits to Institutional Analysis. *Latin American Politics and Society*, 49 (2): 183–200.
- Stephenson, Matthew (2000). A Trojan Horse Behind Chinese Walls?: Problems and Prospects of US-Sponsored ‘Rule of Law’ Reforms in the People’s Republic of China. Harvard. CID Working Paper No. 47, Law and Development Paper No. 3.
- Thomas, Melissa (2007). *What Do the Worldwide Governance Indicators Measure?* August. Washington, DC: SAIS, Johns Hopkins University.
- United Nations Development Programme (2007). *Human Development Report 2007–2008*. New York: UNDP.
- World Bank (2008). *World Development Indicators*. Washington, DC: World Bank. Available at <http://data.worldbank.org/data-catalog/world-development-indicators> [accessed 27 March 2012].

This page intentionally left blank