

Growth and Governance: A Defense*

December 3, 2006

Marcus J. Kurtz
Associate Professor
Department of Political Science
Ohio State University
2140 Derby Hall
154 North Oval Mall
Columbus, OH 43210
614-292-0952
kurtz.61@polisci.osu.edu

Andrew Schrank
Assistant Professor
Department of Sociology
MSC05 3080
1 University of New Mexico
Albuquerque, NM 87131-0001
505-277-1295
schrank@unm.edu

* We'd like to thank Sarah Brooks, Ken Shadlen, and most importantly John Geer for valuable input.

It is clear we have disagreements from our esteemed colleagues on this important issue. The KKM response raises a number of insightful points that advance this debate. We thank them for that and for taking the time to engage our article. As we have reflected on our original essay and KKM's critique of it, we were struck by the fact that we may actually be talking past each other, and that differences in approach between political science and economics may foster this miscommunication. We are not arguing that political science offers a superior approach. To the contrary, our intellectual debt to economics is enormous. Our bibliography offers testimony to that very fact. We nevertheless believe that political scientists have as much to offer economists as vice-versa, including not only a venerable literature on state formation and governance but a no less important tradition of self-conscious concept formation that has been less central to most economists.¹ While political scientists have traditionally been producers of political data, they have, therefore, treated the process of concept formation as a necessary prelude to both measurement and modeling (Sartori 1970; Collier and Mahon 1993; Gerring 1999). By contrast, economists have, until recently, been consumers of political data, and have, therefore, subordinated the need for self-conscious concept formation to the understandable urge to measure and model. We ultimately hope to demonstrate that conceptual issues – and corresponding measurement problems – are at the core of the debate over growth and governance and that progress will be less than ideal until they are addressed. A meeting of minds may be too much to ask for at the present time, but a meeting of method would almost certainly constitute a step in the right direction. We, therefore, offer this rebuttal as not just a response to the thoughtful critique of KKM, but as an effort to bridge these gaps and develop even better ways to make real advances in understanding the relationship between growth and governance.

Concept formation

KKM defined government effectiveness (GE) as the “competence of the bureaucracy and the quality of public service delivery” in their earlier papers (KKM 2005, p. 4). Our doubts about GE derived neither from their definition nor from their conceptual scheme but from the construction of the measure itself. We worried that perception-based indicators of governance in general, and their indicator of GE in particular, suffered from systematic measurement error, selection bias, and halo effects—in short that they did not measure the concept they introduced.

By way of rebuttal, however, KKM have all but abandoned their indicator of GE in favor of distinct indicators (e.g., Rule of Law) designed to capture “a more basic notion of governance going back to the seminal work of Douglas North: the norms of limited government that protect private property from predation by the state” (KKM 2006a, p. 2). We therefore find ourselves in the awkward position of defending their conceptualization of GE in order to criticize their operationalization of GE. Why is an ostensibly narrow focus on the “competence of the bureaucracy” superior to an allegedly encompassing focus on the “protection or private property” (KKM 2006a, p. 8)? We believe that political institutions (e.g., bureaucracy) are conceptually independent of policy decisions (e.g., expropriation) and that the former should be defined and evaluated without regard to the latter. A government that ably protects property, lowers taxes, or privatizes industry is not necessarily more capable than a government that adroitly expropriates property, taxes income, or nationalizes industry—whatever one thinks of the policies in question.

We are by no means the first social scientists to distinguish the enduring rules, norms, or constraints implied by the word “institution” from the at times transitory decisions of powerful public officials. After all, Edward Glaeser and his collaborators have already portrayed estimates of “expropriation risk” derived from commercial data sources employed by KKM as

inadequate indicators of political institutions. “Whatever expropriation risk measures,” they argue, “it is obviously not permanent rules, procedures, or norms supplying checks and balances on the sovereign” (Glaeser et al. 2004, p. 276).

Our point is neither to provoke nor to belabor a “terminological tussle” (KKM 2006a, p. 2) but to acknowledge and underscore the fact that “concept formation stands prior to quantification” (Sartori 1970, p. 1038). While KKM take comfort in the fact that their various indicators are highly correlated with each other, and therefore portray the “protection of private property as a proxy for good governance” (KKM 2006a, p. 8)—if nothing else—in their rebuttal, their optimism is arguably misplaced, for bivariate correlations speak to the reliability, rather than the validity, of political indicators (see, e.g., Munck and Verkuilen 2002, p. 29), and their preferred measure of property protection (i.e., their Rule of Law indicator) is triply problematic in any event. First, they exaggerate the transparency and intelligibility of property claims and thereby add an additional source of error to their indicators. After all, the definition of property is at best controversial and at worst socio-culturally bounded. 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—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. Second, they conflate short-run policy preferences like the protection of private property with enduring institutions like meritocratic bureaucracies and thereby render the assessment of causal effects all but impossible. Are the significant coefficients obtained by KKM—and consumers of their data—attributable to the narrowly institutional features of their preferred indicators or to simultaneously incorporated, and therefore statistically indistinguishable, policy considerations? Unfortunately, the answer is anything but clear. And,

finally, they incorporate policy outcomes (as well as policy preferences) across a wide range of different issue areas and thereby aggravate the already vexing problem of causal inference. In fact, KKM define the Rule of Law as “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” (KKM 2006b, p. 4). Positive RL coefficients are therefore no more interpretable than they are surprising. Do they reflect the impartiality of the police and courts? The benefits of crime control? The returns to social capital and generalized trust? The fact that crime and violence are associated with a host of related social ills? Or the costs of expropriation, tax evasion, informality, patent and copyright infringement, or any of the other transgressions specified in the actual measure? And how would one know?

The literature on concept formation in comparative politics holds that effective concepts discriminate among closely related behaviors, outcomes, or processes and that “for fact-finding purposes it is more profitable to exaggerate in over-differentiation than in over-assimilation” (Sartori 1970, p. 1039). But KKM take an ambivalent stance toward discrimination. On the one hand, they disaggregate “governance” into six different dimensions and, to their inestimable credit, warn consumers of their data not to aggregate their various indicators into an overall index of “good government.” On the other hand, they incorporate institutions, policy preferences, and policy outcomes into the definitions of their preferred indicators and thereby undermine the very benefits of disaggregation.

Furthermore, KKM are growing *less* discriminating over time. Take, for example, GE. While KKM originally defined government effectiveness by way of reference to relatively coherent aspects of bureaucratic organization and behavior, and thereby bracketed the behavior of the

legislature, the sovereign, and the courts, they have recently added “the quality of policy formulation and implementation, and the credibility of the government’s commitment to such policies” (KKM 2006b, p. 4) to their definition, and have thereby brought at least two of the aforementioned three actors, as well as public policies and subjective assessments of their “credibility,” back into the picture. Nevertheless, GE continues to constitute the closest thing KKM have to the sort of purely-institutional measure that political scientists routinely investigate and it therefore offers a “best case” scenario for the KKM indicators more generally.

Measurement and sampling

We maintain that perception-based indicators of governance in general, and the KKM indicator of GE in particular, are overly dependent on the impressions of businesspeople. KKM offer a twofold rebuttal. First, they maintain that they incorporate data not only from businesspeople but from a broader array of sources (e.g., non-governmental organizations, citizen surveys, etc.) into their indicators. And, second, they question the degree to which business and societal preferences diverge in any event by noting the high correlation of business and social surveys.

The former claim is puzzling for several reasons. First, they cite a number of sources (e.g., Reporters without Borders, the US State Department) that are incorporated *not* into their most recent indicator of GE but into distinct measures that are not in dispute. Second, they fail to distinguish “representative” sources that are available for the majority of their countries from “non-representative” sources that are not—and that perforce carry less weight in their overall indicators. Only one of seven so-called representative sources in the most recent release of GE is an NGO, for example, and none is a citizen survey (KKM 2006, Table B3). Six of the seven more widely available sources are commercial risk rating agencies or surveys of businesspeople

that betray the myriad biases we have already described and therefore validate—rather than assuage—our original fears. Third, they exaggerate the neutrality of their noncommercial sources in any event. After all, the Freedom House ratings they use have been portrayed as an “ideologically loaded” product of American values and preferences (Munck 2006, p. 38; see also Munck and Verkuilen 2002). The multilateral development banks are by no means policy neutral (Wade 2002a). And indicators of “electronic governance” assume—rather than prove—that e-governance is good governance and are biased against countries on the other side of the “digital divide” by construction (Wade 2002b). Finally, they fail to acknowledge the full implications of their aggregation scheme. On the contrary, KKM explicitly *assume*—but fail to prove—that the errors of their distinct sources are independent of each other, and therefore employ an aggregation rule that explicitly “rewards conformity” (Kaufmann, Kraay, and Zoido-Lobaton 1999, p. 14). Sources that deviate from common patterns are given less weight in their aggregate indicators than sources that are highly correlated with each other, and the two citizen surveys that weigh so heavily in their response—Afrobarometer and Latinobarometer—are therefore responsible for about 2 percent of their overall rating of GE (KKM 2003, Table 3).² While KKM find that their weighted measure of GE and an “equally weighted” alternative are correlated at $r \geq .97$, and thereby conclude that their weights are immaterial to their estimates (KKM 2006b, p. 25), their alternative measure remains disproportionately dependent upon the perceptions of businesspeople and their advisers as well. After all, GE includes a single citizen survey from each of 36 countries—18 in Latin America and 18 in Africa—and no citizen surveys from Asia, Europe, North America, or Oceania. Consequently, the equally weighted alternative *excludes* the voices of citizens in well over 75 percent of the countries of the world—including eight of the ten largest countries—and relies *disproportionately* upon commercial sources for the

36 countries whose rankings *do* include citizen surveys, since all 36 incorporate ratings from more than one commercial source. Businesspeople and their advisors constitute a tiny—and by most accounts politically distinct—minority of the world’s population but contribute a vastly disproportionate share of the GE index under *either* weighting scheme and we are therefore neither surprised nor convinced by the high reported correlation—which is a test of reliability rather than validity in any event.

Of course, KKM’s biased sample is immaterial if errors are really independent and perceptions of GE are impervious to the respondent’s occupation or position. What, then, are we to make of their Table 1? Do the reported correlations prove that firms and citizens perceive their public sectors in the same way? We preface our answers to both question with a by now redundant caveat—that bivariate correlations are in and of themselves tests of reliability rather than validity—and a benchmark drawn from the expansive literature on the conceptualization and measurement of democracy. The various measures of democracy designed and used by political scientists are correlated at $r \geq .8$ (Przeworski et al. 2000, pp. 56-57). While the Afrobarometer survey and the GCS data invoked by KKM betray a reasonably high correlation ($r = .7$), the former is deliberately weighted toward 18 atypically liberal African regimes, and the more comprehensive Latin American data that go all but unremarked upon in their text diverge markedly. In fact, the correlation between the Latinobarometer survey of citizens *throughout* noncommunist Latin America and the GCS survey of firms—allegedly measures of the *same* concept—is a modest .52, and the mean correlation between the Latin American data and the six commercial risk assessments found in their Table 1 is a mere .42.³

Nor is Latinobarometer exceptional. The mean correlation for the eight commercial risk assessments and firm surveys incorporated into their Table 1 is .75. The mean correlation for the

eight noncommercial sources (e.g., public agencies, NGOs, citizen surveys) that are both included in the table and incorporated into their most recent GE indicator is .62. And the difference in means is significant at $p \leq .08$.⁴

Nevertheless, the Afrobarometer data permit a simpler test of KKM's faith in the insignificance of the differences between business and popular perceptions of GE.

Afrobarometer includes data on whether the respondent is a "businessperson." Table 1 includes logistic regressions of individual responses to all of the Afrobarometer questions included in the most recent GE indicator on an indicator variable coded 1 if the respondent is a businessperson and a series of country dummies. Positive assessments are coded 1; businesspeople are coded 1; odds ratios are placed next to their parenthesized p values; and the invariably significant results suggest—contra KKM—that businesspeople have *better access* to government services and nonetheless hold their governments in *lower regard* than their compatriots. What are the implications for GE? The fact that businesspeople and citizens within the same country part company on the question of good governance speaks to a deeper issue that underpins the entire debate: We don't even have common perceptions of government effectiveness at the national level let alone the ability to construct cross-nationally valid rankings.⁵

Furthermore, the Afrobarometer data present a relatively conservative test of our claim. After all, the survey questions cover more or less uncontroversial issues like household services, education, and health care provision. The risk-rating authorities who are responsible for the bulk of the GE indicator address more controversial topics like taxation, regulation, and the existence and interpretation of red tape—where the interests of firms and citizens are particularly likely to diverge.

In fact, KKM's colleagues at the World Bank devote an entire subsection of a recent *World Development Report* to the "basic tension" between "firm preferences" for limited taxation, regulation, and spending and the overarching "public interest" (World Bank 2004, p. 37). Why and to what effect would KKM depart from the position taken by their employer's "flagship publication" (Wade 2002a, p. 220)? We worry that KKM are implicitly nullifying a decade of *glasnost* in the international donor community. After all, the multilateral development banks no longer deny industrial policy's potential contribution to growth and development; instead, they hold that industrial policy demands better governance than the typical less developed country (LDC) can muster, and therefore portray free market reform as a second best alternative. By treating *laissez faire* as part of the very definition of government effectiveness, however, KKM and their adherents condemn industrial policy by tautology: Where does industrial policy work? Where governments are effective. What defines effective government? The absence of red tape and regulation—that is, the absence of the very lifeblood of industrial policy.

Halo effects revisited

We worried that perception-based measures of governance were also contaminated by halo effects. KKM responded by demonstrating that our statistical results were vulnerable to defensible—if by no means unassailable—changes in specification and that GE responds not to "recent growth performance" but to "long-run growth in the 20 years prior to the date of the governance indicator" (KKM 2006a, p. 6). We have neither the space nor the inclination to debate the relative merits or interpretation of different model specifications at length (however see Kurtz and Schrank 2006), and would rather take the opportunity to make two simple but by no means unimportant points about concept formation.

First, we need not win the debate over halo effects to win the debate over GE. We have already established that GE (1) conflates policy preferences and outcomes with political institutions; and (2) relies almost entirely upon the perceptions of businesspeople and their advisers in any event. Consequently, the correlation between GE and growth—if any—is at best uninterpretable *whether halo effects are demonstrable or not*. A positive GE coefficient could reflect any combination of at least three different underlying processes: first, a positive relationship between the *policies* captured by the measure (e.g., deregulation, tax relief, school spending, etc.) and growth; second, a positive relationship between *actual government efficacy* (e.g., bureaucratic capability) and growth; or, third, *herd behavior* 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. We need not win the debate over halo effects, therefore, to win the debate over GE, for conceptualization and measurement are, as political scientists have long noted, logically prior to hypothesis testing. We need only demonstrate that their indicator of GE relies disproportionately upon the perceptions of businesspeople *or* conflates policy preferences and outcomes with political institutions. We do both.

Second, the models presented in Table 2 cast at least as much doubt on the validity of GE as they do upon the likelihood of halo effects. Why? KKM depart from past practice, including their own past practice (see, e.g., Isham and Kaufmann 1999; Dollar and Kraay 2003), by abandoning our direct measure of the stock of human resources for a “more widely available” indicator of their flow: the gross secondary enrollment rate. But the gross secondary enrollment rate is more accurately portrayed as a *reflection* than a *cause* of government effectiveness—especially in light of their operationalization of GE. After all, KKM incorporate “the quality of public service delivery” into the very definition of GE (KKM 2005, p. 4), and their reconstituted

models therefore treat a *subjective* measure of service delivery (i.e., GE) as a function of an *objective* measure of service delivery (i.e., enrollment rates). We are not the first social scientists to realize that secondary enrollment constitutes a measure, rather than a proximate cause, of government effectiveness (Migdal 1988, p. 286), and we are therefore puzzled by the objective measure's persistent inability to predict the subjective measure with confidence.⁶ Are the insignificant coefficients on the enrollment variables in the models distilled into their Table 2 products of specification error on the right-hand side or the outright invalidity of the measure on the left-hand side? We fear the former. We suspect the latter. And we note that neither interpretation bodes well for KKM.

The Validity and Virtue of Prospective Models

Our models of prospective growth are designed to mitigate the problem of endogeneity as well as the consequences of halo effects. While we believe that endogeneity is less vexing and halo effects are less likely in prospective models, and therefore regress growth rates on prior measures of GE in Tables 2-4 of our original paper, KKM assert "that the growth effects of good institutions show up only over time" and therefore call the very basis "for looking at the relationship between very noisy short-run fluctuations and governance" into question (KKM, pp. 7-8). We are well aware of the noisy nature of short-run growth rates but are nonetheless surprised by KKM's willingness to admit that GE is unlikely to be perceptibly related to prospective short-run growth on average—especially since *their* measure incorporates *several* questions about anticipated *short-run* growth rates (KKM 2006, Table B3) and thereby (1) provides the allegedly absent "justification" for examining prospective growth over short periods and (2) is biased in favor of a *positive* finding from the outset.⁷

Why, then, does their faith in their measure survive the appearance of a *negative* finding? Shouldn't the apparent absence of a correlation between governance and short-run growth call their conviction into question? We believe the results should at least cast doubt upon the consensus and open the door to conceptual and causal reflection. But KKM give the benefit of *any* doubt to students of growth and governance over the *very* long run who *at a minimum*: (1) make heroic assumptions about the worldwide distribution of income circa 1800; (2) conflate the risk of expropriation (i.e., a policy) with the effectiveness of government (i.e., an institution); and (3) instrument for the risk of expropriation in the late twentieth century with the mortality rates of the most violent expropriators in world history—European colonists—in earlier centuries and thereby implicitly illustrate—but fail to take account of—the socio-culturally bounded nature of property claims. We've already discussed the limits to the by no means unassailable (see, e.g., Glaeser et al. 2004; Bardhan 2005) existing literature in detail (see, e.g., KS, notes 7 and 10) and see no need to revisit the issue other than to note that by invoking a consensus to refute a challenge to the consensus, KKM are pioneering a new and to our minds intolerably conservative approach to social scientific argumentation.

Conclusion

We thank KKM for drawing renewed attention the problem of governance in developing countries, and for opening the door to continued and productive interdisciplinary dialogue on the matter. Political sociologists from Max Weber onward have drawn a distinction between the institutional “rules of the game” and the policies produced by the players. While KKM accept the distinction in theory, and portray their estimates as indicators of institutions, they (1) elide the distinction in practice by incorporating policy preferences and outcomes into their indicators

and simultaneously (2) compound their error by using a biased sample of sources. The results are triply troubling. First, KKM render the interpretation of their indicator all but impossible. Do positive GE coefficients—if any—reflect bureaucratic capacity, the influence of policy preferences (or outcomes) that are implicitly or explicitly incorporated into their measures (e.g., letting foreign firms “go about their business”), halo effects, or the potentially irrational exuberance of ill-informed investors? And how would one know? Second, KKM send LDC policymakers mixed signals. After all, the MDBs take pride in their widely-acknowledged influence over the LDCs (see KS, p. 23) and the United States government is already conditioning aid allocations on performance standards established in part by KKM. The intuition is straightforward: LDCs should be rewarded for good behavior and punished for the opposite. But what is good behavior? What policies will foster an improved GE score? And will they foster growth and development as well? Unfortunately, the answers are anything but clear, for the sources employed by KKM reward aid recipients for policies that are almost certainly in tension with each other including stabilizing their polities, deregulating their markets, lowering their tax rates, ensuring the health and well-being of their citizens, maintaining macroeconomic stability, providing reliable infrastructure, and guaranteeing the skill and integrity of their civil servants (KKM 2006, Table B3). What, then, are aid recipients to do? Raise taxes so as to provide health care and education? Place social and political stability at risk by cutting spending? Add to the social service burden by liberalizing prices? Almost every potential solution aggravates another problem, and the KKM benchmarks thereby punish poor countries for their very poverty. If they could solve their social and economic problems, after all, they wouldn’t need foreign aid in the first place. Third, and finally, KKM all but ignore the entreaties of economists who realize that “econometric work should be informed by detailed studies of

individual countries” (Durlauf et al. 2005, p. 646). We have gone to great lengths to integrate qualitative as well as quantitative material and have therefore invoked the experiences of a number of developing countries that have fallen short of the KKM benchmarks and nonetheless performed well in macroeconomic terms (e.g., Korea, the Dominican Republic, postwar Italy, etc.). KKM bypass the case study literature and instead compile annual data on “institutional” factors that they simultaneously portray as constant over the course of centuries. We find their oversight unfortunate. We find their data perplexing. And we can’t help but think that their evidence and findings are at least in part products of their failure to fully engage the case study literature.

Political scientists have long believed that debates over concepts are at least as fertile as tests of causal inference. David Laitin illustrates the point by noting that “concepts such as ‘charisma’ and ‘the division of labor’ have been longer-lasting than any valid claims about the causal effects of these concepts” (Laitin 1995, p. 455). And conceptual innovation is not only prior to quantification but is frequently a product of qualitative research. We therefore conclude not only by underscoring the importance of conceptual rigor but by affirming the value—if by no means primacy or exclusivity—of qualitative research and, finally, by acknowledging political science’s unique contributions on both fronts.

References

- Ball, Terence, James Farr, and Russell Hanson. 1989. *Political Innovation and Conceptual Change*. Cambridge: Cambridge University Press.
- Bardhan, Pranab. 2005. "Theory or Empirics in Development Economics?" *Economic and Political Weekly* 40 (40).
- Collier, David, and James Mahon. 1993. "Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Analysis." *American Political Science Review* 87 (4): 845-55.
- Dollar, David, and Aart Kraay. 2002. "Growth Is Good for the Poor." *Journal of Economic Growth* 7(3): 195-225.
- Durlauf, Steven, et al. 2005. "Growth Econometrics." In *Handbook of Economic Growth*, eds. Philippe Aghion and S. Durlauf. Amsterdam: North Holland, 555-677.
- Gerring, John. 1999. "What Makes a Concept Good? A Criterial Framework for Understanding Concept Formation in the Social Sciences." *Polity* 31(3): 357-93.
- Glaeser, Edward, et al. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3): 271-303.
- Isham, Jonathan, and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *Quarterly Journal of Economics* 114(1): 149-84.
- Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobaton. 1999. *Aggregating Governance Indicators*. Typescript.
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2003. *Governance Matters III: Governance Indicators for 1996-2002*. Typescript.
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2005. *Governance Matters IV: Aggregate and Individual Governance Indicators for 1996-2004*. Typescript.

Kurtz and Schrank

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2006. "Growth and Governance: A Response." *Journal of Politics* 69(2).

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2006. *Governance Matters V: Aggregate and Individual Governance Indicators for 1996-2005*.

<<http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contentMDK:21045419~menuPK:1976990~pagePK:64168445~piPK:64168309~theSitePK:1740530,00.html>> December 3, 2006.

Kurtz, Marcus, and Andrew Schrank. 2006. "Growth and Governance: Statistical Questions." To be available at <http://psweb.sbs.ohio-state.edu/faculty/mkurtz/papers.htm>.

Laitin, David. 1995. "Disciplining Political Science." *American Political Science Review* 89 (2): 454-6.

Migdal, Joel. 1988. *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton: Princeton University Press.

Munck, Gerardo, and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35 (1): 5-34.

Munck, Gerardo. Forthcoming. "Adam Przeworski: Capitalism, Democracy, and Science." In *Passion, Craft, and Method in Comparative Politics*, eds. Gerardo Munck and Richard Snyder. Baltimore: Johns Hopkins University Press.

Przeworski, Adam, Michael Alvarez, José Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge: Cambridge University Press.

Sartori, Giovanni. 1970. "Concept Misformation in Comparative Politics." *American Political Science Review* 64(4): 1033-53.

Kurtz and Schrank

Wade, Robert. 2002a. "US hegemony and the World Bank: the fight over people and ideas."

Review of International Political Economy 9(2): 215-43.

Wade, Robert. 2002b. "Bridging the Digital Divide: New Route to Development or New Form

of Dependency?" *Global Governance* 8(4): 443-67.

World Bank. 2004. *World Development Report 2005: A Better Investment Climate for*

Everyone. Washington, D.C.: World Bank.

Table 1: Do businesspeople and their neighbors perceive government in the same way?

<i>Question</i>	<i>Responses</i>	<i>Businessperson</i>
Based on your experiences, how easy or hard is it to obtain household services (like electricity or telephone)?	0 = never try, very difficult, or difficult;	1.755 (p = .000)
Based on your experiences, how easy or hard is it to obtain identity documents (like birth certificate, passport)?	1 = easy or very Easy	1.227 (p = .001)
How well or badly would you say the current government is handling health care?	0 = very badly or fairly badly; 1 =	.889 (p = .053)
How well or badly would you say the current government is handling education?	fairly well or very well	.777 (p = .000)

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. We have dichotomized the four and five point Likert scales of answers for ease of calculation and interpretation and dropped respondents who were unwilling or unable to answer. The data are from Round 2 of Afrobarometer (2002-3); the more recent data available to KKM are not publicly available.

Endnotes

¹ An admittedly unscientific review of graduate methodology syllabi in the top five political science and economics departments confirms our suspicions. Political science departments almost invariably include at least one course with a substantial component on concept formation and measurement. Economics departments no less consistently focus on econometrics and modeling. For a theoretical analysis of important political science concepts, see, e.g., Ball et al. (1989).

² KKM released weights that would be applied “to a hypothetical country appearing in all of the available sources for that indicator” in 2002 (KKM 2003, p. 45).

³ Only two of six correlations between Latinobarometer and commercial risk assessments are significant at $p \leq .05$ and two more at $p \leq .10$. The correlations confirm the reliability but not the validity of the measures in any event. Commercial and non-commercial sources could diverge markedly on average and nonetheless correlate highly.

⁴ The inclusion of the Columbia University data would drop the p value to .11; however, the Columbia data have been dropped from the most recent GE indicator and their inclusion in Table 1 is therefore puzzling.

⁵ We couldn’t even use the KKM data to construct ordinal rankings without assuming that the direction and degree of perceptual differences between citizens and businesspeople are constant across countries and uniform across indicators. The former is unlikely and the latter is demonstrably untrue in the Afrobarometer data—as well as unknowable for the bulk of the world’s population in light of the paucity of citizen surveys in their data set..

⁶ KKM present only the coefficients on the halo effects. We therefore replicated their models with their data and found that the *minimum* p value for secondary enrollment is 0.343.

⁷ KKM briefly discuss the “difficulty of estimating growth regressions” in footnote 12 and ask why we abandon the random effects estimators used in Table 1 for the fixed effects models in Tables 3 and 4. We note that the dependent variables in Table 1 and Tables 3-4 differ; the fixed effects in the growth models are standard tools to address the problem of unobserved country-level heterogeneity; and the—by no means uncontroversial (see Durlauf et al. 2005)—alternative estimation procedures recommended by Caselli et al. in the article cited by KKM appear to generate consistent results.