Growth and Governance: A Reply

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi
The World Bank[1]
September 2006

In this issue of the Journal of Politics, Marcus Kurtz and Andrew Shrank (hereafter KS) offer a sweeping critique of the exisiting literature on governance and growth. They argue that perceptions-based cross-country measures of governance, and in particular those we have constructed in our ongoing work[i], are fatally flawed.   They also produce empirical evidence which they claim shows that perceptions of governance are driven by short-term growth performance. Finally they argue that there is little convincing evidence that good governance spurs growth.

In our response we show that these claims are unsubstantiated. We first show that their claims of biases in perceptions-based measures of governance are speculative and, to the extent that they are falsifiable, do not withstand empirical scrutiny. We next show that the empirical evidence in support of their claim that governance perceptions respond to short-run growth is both statistically fragile and conceptually flawed. Finally we dismiss their empirical work on the effects of governance on growth, which we argue is far removed from the best-practice frontier in cross-country growth empirics. We instead briefly describe the rich body of recent work in the economics literature that has documented a sizeable long-run effect of governance on growth.

[1] 1818 H Street NW, Washington DC 20433, dkaufmann@worldbank.org, akraay@worldbank.org, mmastruzzi@worldbank.org. The opinions expressed here are the authors' and do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent.

Getting Concepts Straight

Before delving into the details of our response, we first note that KS's definition of governance is in our view convoluted and ultimately too narrow. In their opening sentence they refer broadly to "political corruption and malgovernance". By the next paragraph they have shifted to a much narrower concept of the "quality of public administration". By page 11 the definition has shifted again to "the ability of the state to formulate and implement its goals". KS emerge from these shifting definitional sands to focus on just one of the six measures of governance that we construct, "Government Effectiveness".

While it is easy to get into endless terminological tussles over what governance is, here we simply want to make the point that KS focus narrowly on just one of our specific measures of governance and in our view inappropriately ignore other dimensions of governance that have received much more attention in the empirical literature on governance and growth. As we discuss further below, leading papers in this literature tend to focus on 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. This concept is much more closely related to our measures of Rule of Law and Control of Corruption, as well as several other indicators of these concepts.

We do not want to make too big a deal of this conceptual distinction because in the end these aspects of governance tend to be quite highly correlated across countries. There are not many countries where corruption is high yet the public sector manages to provide public services effectively, which is what our Government Effectiveness measure captures. Yet for the sake of conceptual clarity, and for the sake of placing this paper in the context of the existing empirical literature, it is important to note that KS's focus on this one specific measure of governance is very narrow indeed.

He Says, She Says: Are Perceptions Measures Really Biased?

KS begin their discussion with a series of assertions that perceptions-based measures of governance are biased in various ways. KS first argue that businesspeople's views of what good governance is might be very different from other views more broadly reflective of the public interest. In short, they argue, businesspeople like low taxes and less regulation, while the public good demands reasonable taxation and appropriate regulation. Estimates of governance based on the perceptions of businesspeople will therefore necessarily be biased.[ii]

In response, we note first that our six aggregate governance indicators rely on much more than just the views of businesspeople. In the latest 2005 update of our governance indicators, our data sources include four cross-country surveys of firms, as well as seven commercial risk rating agencies, which one might think reflect narrower business interests. But we also rely on three cross-country surveys of individuals, six sets

of ratings produced by government and multilateral organizations (such as the World

Bank, the African Development Bank and the US State Department), and finally another

11 data sources produced by a wide range of non-governmental organizations (such as

Freedom House, Reporters Without Borders, and many others). It is therefore simply

incorrect to dismiss our indicators as reflecting solely the narrow interests of the business

elite.

While we accept that anti-regulation biases in surveys of businesspeople are

possible in principle, the more relevant question is whether such biases are practically

important. Here KS offer us no empirical evidence. This is unfortunate because the

various hypotheses of bias that they advance lend themselves well to empirical testing.

Consider for example the argument that businesspeople have a view of what constitutes

good governance that is fundamentally different from other types of respondents. If this

is true, then the responses of firms (or commercial risk rating agencies who serve mostly

business clients) to questions about governance should not be very correlated with ratings

provided respondents who are more likely to sympathize with the common good, such as

individuals, NGOs, or public sector organizations.

This turns out not to be the case. Table 1 reports some simple correlations of

assessments of Government Effectiveness from one of our cross-country surveys of

firms, the Global Competitiveness Survey (GCS), with all of our other data sources for

this measure in 2005.[iii] We first consider two other surveys of firms, the World

Competitiveness Yearbook and the Business Environment and Enterprise Performance

Survey (BEEPS). If businesspeople have a monolithic view of what constitutes good governance we should expect these two surveys of firms to be highly correlated with the World Competitiveness Yearbook. This is not obviously the case, with correlations of 0.74 and 0.43 respectively. In fact it is notable that the correlation with the BEEPS is lowest among all sources reported in this table. While it is true that the firm survey responses tend to be fairly highly correlated with commercial risk rating agencies, they are also quite highly correlated other data sources. It is striking for example that the correlation between firm responses and individual responses from the Afrobarometer survey is 0.7 across the 23 countries in Africa covered by these two surveys. It is also striking that the correlation of the GCS with the World Bank and African Development Bank's assessments of policy and institutional quality is around 0.75. While it is true that there are some data sources that are not very highly correlated with the GCS, we do not think one could reasonably conclude from Table 1 that there is a substantial bias in the responses of businesspeople relative to those of other types of respondents. More systematically, in Kaufmann, Kraay and Mastruzzi (2006b) we document how the country rankings provided by our aggregate indicators are very robust to alternative weighting schemes. This can only be the case if on average there is substantial consensus among our different data sources regarding the broad concepts of governance being measured.

KS also argue that cross-country surveys of both firms and individuals are affected by cultural biases. For example, respondents in different countries might have different norms as to what does or does not constitute corruption. Presumably however

these cultural biases should not be present in cross-country expert assessments, that are deliberately designed to be comparable across countries. This in turn suggests that low correlations between surveys and expert assessments would be consistent with cultural biases. Yet what is striking is that surveys of firms tend to be quite highly correlated with expert assessments of all types. This can be seen in Table 6 of Kaufmann, Kraay and Mastruzzi (2006b) where we document sizeable correlations between expert assessments and the GCS, for all six of our dimensions of governance.

We have also in our past work looked at other potential sources of bias in our data sources, that are not mentioned by KS. One concern often heard is that the ratings provided by NGOs and think-tanks tend to be coloured by the ideological orientation of the organization providing the ratings. In Kaufmann, Kraay, and Mastruzzi (2004) we devised a simple test for such political biases. We examined whether the difference between the assessments of think-tanks and firm surveys was systematically correlated with the political orientation of the government in power in the countries being rated. We found that this was generally not the case, casting doubt on this additional possible source of bias.

Even so, we would not want to argue that either cross-country surveys or expert assessments of various sorts are perfect measures of governance. We have long acknowledged that every one of our underlying data sources -- or any other potential measure for that matter -- is at best an imperfect proxy for governance. This is precisely why we think it is useful to construct aggregate governance indicators that combine

information from many different sources: in doing so we are able to smooth out some of this measurement error, with the result that the aggregate indicators are more reliable measures of governance than any of our individual indicators.

Growth and Perceptions of Governance: How Shiny is Your Halo?

KS next argue that perceptions-based measures of governance respond to recent economic performance and thus do not reliably capture deeper notions of institutional quality. This concern, which is sometimes referred to as the "halo effect", is also not new and has been applied to various indicators of governance.[iv] KS go on to provide some empirical work which they claim provides evidence of this sort of bias. They estimate cross-country regressions of our Government Effectiveness measure on per capita GDP, a measure of human capital, the logarithm of country population, and per capita GDP growth in the two previous years. The first three variables are intended as controls for fundamental determinants of governance, and the growth variable is intended to pick up "halo" effects. They find that prior growth is significantly correlated with better governance ratings. Based on this they argue that governance indicators based on perceptions data are unreliable because they reflect recent short-run economic performance rather than more fundamental cross-country differences in institutional quality.

We do not find this exercise compelling for several reasons. First, the finding is simply not robust to very minor but reasonable changes in their empirical specification.

Kaufmann, Kraay, and Mastruzzi

We show this in Table 2. In the first panel we summarize a set of regressions that essentially replicate KS's Table 1. In particular we replicate the first four cross-sectional regressions from that Table, with Government Effectiveness as the dependent variable.[v] We estimate these regressions for 1996, 1998, 2000 and 2002 as do KS, and also for 2003, 2004, and 2005 using the most recently-available update of our governance indicators. Moreover, we think that KS's halo effects critique could in principle apply to all six of our aggregate governance indicators, so we also estimate the same regressions for our other five indicators. This results in a set of 42 regressions. To conserve space we report only the coefficient on lagged growth, so that each cell in the table refers to a separate regression. Consistent with KS we find that for our Government Effectiveness measure, lagged growth is significantly positively correlated with the governance indicator (third column of top panel of Table 2). This is true not only for the years 1996-2002 considered by KS, but also for 2003-2005. Moreover, the other columns show that in many cases, lagged growth is significantly correlated with our other governance indicators: in 32 of the 42 regressions we find a significant positive correlation at the 5 percent significance level.

In the next panel we show the effect of two minor departures from the original KS specification. Instead of entering per capita GDP in levels as they do, we enter it in log-levels. This is very standard practice in cross-country empirics, and statistically is more appropriate since the relationship between the dependent variable and log per capita GDP is much closer to being linear and we are using a linear regression model. In addition, we note that KS's results cover a relatively small set of around 100 countries. The key

constraint here is the limited availability of the human capital variable that they use -- the governance indicators are available for over 200 countries. We therefore also expand the sample of countries by replacing the stock of human capital variable with a more widely-available flow measure: the gross secondary school enrollment rate.[vi]

The new regressions in this panel cover a much larger sample of around 150 countries. Now the results are drastically different from before, in that we find only six of 42 cases where lagged growth enters significantly and positively. And more peculiarly, there are seven cases where lagged growth enters significantly *negatively*, which is just the opposite of what one would expect if halo effects are important. In the remaining 29 cases there is no significant correlation between our governance measures and growth in the previous two years. In fact, looking only at the sign of the estimated coefficients, we find that they are nearly evenly split between 24 positive and 18 negative. The fact that these two minor -- but we think very reasonable -- variants eliminate the significance of KS's results suggest to us that their findings are simply not robust. Or put differently, in order for KS's critique to be convincing they would need to also provide an account of why halo effects are important in their particular 100 country sample but not in a broader 150 country sample.

Nevertheless, suppose we take the KS sample and specification at face value, despite its lack of robustness. We next argue that KS are mistaken to interpret the significance of the coefficient on lagged growth as evidence of halo effects, because we might very well observe this partial correlation in the data even if halo effects were not

present in the data. The reason is simple. Suppose, reasonably enough, that countries with effective governments do grow faster, at least in the long run. Suppose, also reasonably, that government effectiveness is fairly persistent over time. Then we should expect to see a correlation between long-run growth, for example, growth over the previous 10 or 20 years, and our measures of Government Effectiveness. This would not be due to halo effects, but rather would simply reflect the beneficial growth effects of having a competent government over a long period of time. However, to the extent that recent growth performance is correlated with long-run growth performance, we might very well find that the former is in fact correlated with our measure of Government Effectiveness simply because the latter is omitted from the regression.

This omitted variable problem turns out to be important in KS's preferred specification. The easiest way to see this is to again replicate the regressions in KS's Table 1, but now adding a variable capturing long-run growth in the 20 years prior to the date of the governance indicator. We do this in Table 3, for KS's basic specification for Government Effectiveness. In all of the five periods shown, we find that prior 20-year average growth is significantly correlated with Government Effectiveness. Moreover, now in only one case is recent growth performance significantly correlated with government effectivness.

One could nevertheless still argue that the significance of prior long-run growth in these regressions constitutes evidence of halo effects, with subjective perceptions of government effectiveness being tainted by previous long-run economic performance as

well. However, we do not think this would be credible. To show this, we also estimate the same regression, but instead replace the dependent variable with the Evans and Rauch (1999) measure of professionalism in the civil service, which KS laud as a carefully-constructed measure of institutional quality likely to be free of any perceptions biases (although unfortunately covering only a small sample of 35 countries at one point in time). Again we find that prior long-run growth performance is significantly correlated with this measure of institutional quality which is unlikely to be tainted by any kind of halo effects. And once we control for long-run growth, short-run growth has no significant correlation with this measure of institutional quality. This supports our interpretation of the data: government effectiveness is both persistent over time, and correlated with long-run growth, which in turn is correlated with short-run growth. This explains the spurious significance of short-run growth in the KS regressions.

Suppose despite all this one were to insist on interpreting the significance of short-run growth in KS's specifications as evidence of halo effects. A final question one might ask is whether these effects practically important or not, in the sense of significantly contributing to the variation in our governance indicators. We answer this question by calculating the share of the cross-country variation in our Government Effectiveness measure that is due to the estimated halo effect, defined as prior growth multiplied by its estimated coefficient.[vii] We find that this share is quite small, ranging from a low of 1.6 percent in 2004 to a high of 5.6 percent in 2000. While we have argued at length that the KS's results are not robust and their interpretation is flawed, even if we

Kaufmann, Kraay, and Mastruzzi

take them at face value, the halo effects they claim to document do not appear to be quantitatively large.

Governance and Growth: Putting the Straw Man Out of His Misery

In the final part of their paper KS investigate the relationship between our measure of Government Effectiveness and subsequent growth, in a large cross-section of countries. The left-hand-side variable in their regression is real per capita GDP growth in the two years following the date of the governance indicator. The right-hand-side variables consist of our Government Effectiveness measure and a set of other control variables. They make much of the fact that, conditional on these other variables, good scores on the Goverment Effectiveness indicator are not correlated with faster subsequent growth. In fact, they conclude from this that there is no evidence that good governance raises growth performance in the next two years.

We think this final exercise is little more than a straw man, for several reasons. To begin, we note that KS appear to completely bypass a large and careful literature that has studied the connection between institutional quality and long-run economic performance. For example, in a seminal empirical paper Knack and Keefer (1995) documented a highly significant partial correlation between various measures of institutional quality and 25-year average growth rates across countries, controlling for a variety of factors. In another highly-influential paper, Acemoglu, Johnson and Robinson (2001) have shown that the historically-determined component of institutional quality has

had a strong causal effect on current levels of per capita income across countries today.

Since cross-country differences in per capita incomes today primarily reflect differences

in these countries' very long-run growth rates, this paper can also be interpreted as

capturing a relationship between institutional quality and subsequent very long run

growth. These two papers, as well as many others in this literature, have helped to shape

the views of many academics and policymakers that good governance is important for

growth.[viii]

Should this view be shaken by the couple of regressions that KS offer? We think

not. First, we note that KS's choice of very short-run growth as the dependent variable

differs from most of the existing literature which has focused on long-run growth. We

think this latter focus is much more appropriate. It seems to us quite plausible that the

growth effects of good institutions show up only over time, while short-run fluctuations

unrelated to institutional quality are likely to dominate a lot of the variation in yearly

growth rates. It is for this reason that virtually the entire empirical cross-country growth

literature has focused on long-run average growth rates. In contrast, KS offer no

justification for looking at the relationship between very noisy short-run fluctuations and

governance.

Second, we have already noted that the literature on institutions and growth has --

for good reasons-- mostly focused on more fundamental notions of protection of private

property as a proxy for good governance. For example, the Acemoglu, Johnson and

Robinson (2001) paper emphasizes the importance of property rights protection as

proxied by a measure of expropriation risk. We recognize that various measures of property rights protection tend to be fairly correlated with our Government Effectiveness measure that KS use, and it is likely that KS would obtain similar results in their problematic specification if they used these other measures. However, we think that KS could do a much better job of justifying their focus on just this one very particular measure of governance to the exclusion of arguably more fundamental ones studied by the existing literature.

Third, KS depart from most of the recent literature on institutions and growth by sidestepping entirely the fundamental question of causality -- do observed correlations between governance and growth reflect causation from governance to growth or the other way around?[ix] Or do they reflect the effect of a myriad of potential other variables not included in the regression that drive both growth and governance. This econometric difficulty is by now very well-understood, and leading papers in the literature have come up with a variety of creative strategies for sorting out the causal effect of good governance.[x] The strategy of naively estimating cross-country regressions by ordinary least squares alone as is done by KS has long been abandoned by the serious empirical growth literature.[xi]

Finally, we note that the regressions presented by KS contain some quite implausible estimates of the effects of other growth determinants. Should we take seriously their findings on governance and growth when the very same regressions also tell us that macroeconomic stability (as captured by inflation) is insignificantly correlated

with growth, whereas the empirical growth literature has showed that at least very high inflation rates are correlated with slower growth (columns 1-4 of their Table 3)? Should we take seriously their finding in column 4 of Table 3 that higher investment rates are significantly *negatively* correlated with growth? If KS would like us to take seriously the lack of significance of governance in their growth regression, they should also provide an account for these other quite peculiar findings.

In conclusion, we do not wish to argue that the empirical literature on governance and growth is now conclusive so that no further work is needed -- to do so would be complacent in the extreme. We are acutely aware of the limitations of exisiting cross-country measures of governance, and have long argued for the need to complement cross-country indicators with more detailed and nuanced within-country data in order to inform efforts to improve governance at the country level. There is also plenty of room for further serious work in understanding both the causes and consequences of good governance, at the cross-country and within-country level. We do however think that contributions to this exciting and important research agenda would do better to take the large existing literature as a starting point, and then document whether these findings can be overturned -- or refined -- in reasonable ways. Starting *de novo* with a few flawed growth regressions as KS do seems to us unlikely to be helpful in advancing our understanding of these important issues.

References

Acemoglu, Daron, Simon Johnson, and James Robinson. "The Colonial Origins of Comparative Development. *American Economic Review*. 91(5):1369-1401.

Evans, Peter and James Rauch (1999). "Bureaucracy and Growth: A Cross-National Analysis of the Effects of 'Weberian' State Structures on Economic Growth". *American Sociological Review* 64 (October 1999): 748-765.

Chong, Alberto and Cesar Calderon (2000). "Causality and Feedback Between Institutional Measures and Economic Growth". *Economics and Politics.* 12(1):69-81.

Glaeser, Edward, Rafael LaPorta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2004). "Do Institutions Cause Growth?". Journal of Economic Growth. 9(3):271-303.

Hall, Robert E., and Charles Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114(1):83-116.

Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobatón (1999a). "Aggregating Governance Indicators." World Bank Policy Research Working Paper No. 2195, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/aggindicators.html.

——— (1999b). "Governance Matters." World Bank Policy Research Working Paper No. 2196, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/govmatters.html.

——— (2002). "Governance Matters II – Updated Indicators for 2000/01." World Bank Policy Research Working Paper No. 2772, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/govmatters3.html.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2004). "Governance Matters III: Governance Indicators for 1996, 1998, 2000, and 2002". World Bank Economic Review. 18:253-287.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2006a). "Measuring Governance Using Perceptions Data" in Susan Rose-Ackerman, ed. *The Handbook of Economic Corruption*. Edward Elgar.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2006b). "Governance Matters V: Governance Indicators for 1996-2005. World Bank Policy Research Department Working Paper No. 4012.

Kaufmann, Kraay, and Mastruzzi

Klenow, Peter and Andres Rodriguez-Clare (1997). "The Neoclassical Revolution in
    Growth Economics: Has it Gone too Far?", in Ben Bernanke and Julio
    Rotemberg, eds. NBER Macroeconomics Annual. Cambridge, MIT Press.

Knack, Steven and Philip Keefer (1995). "Institutions and Economic Performance:
    Cross-Country Tests Using Alternative Measures." *Economics and Politics*, 7,
    207-227.

Rigobon, Roberto and Dani Rodrik (2004). "Rule of Law, Democracy, Openness, and
    Income: Estimating the Interrelationships. Manuscript. MIT and Kennedy
    School

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi (2004). "Institutions Rule:
    The Primacy of Institutions over Geography and Integration in Economic
    Development". *Journal of Economic Growth* 9(2):131-165.

Shleifer, Andrei and Robert Vishny (1999). *The Grabbing Hand: Government
    Pathologies and their Cures*. Cambridge: Harvard University Press.

Table 1: Government Effectiveness: Comparing View of Different Respondents

Table reports the correlation of the 2005 assessment of Government Effectiveness from the Global Competitiveness Report's survey of firms with other sources

Surveys of Firms	
World Competitiveness Yearbook	0.74
BEEPS	0.43
Commercial Risk Rating Agencies	
Economist Intelligence Unit	0.86
Political Risk Services	0.75
Global Insight DRI	0.76
World Markets Online	0.88
Merchant International Group	0.73
Business Environment Risk Intelligence	0.87
Public Sector Agencies	
World Bank CPIA Ratings	0.74
African Development Bank CPIA Ratings	0.75
Asian Development Bank CPIA Ratings	0.42
NGOs	
Columbia University State Capacity Project*	0.79
Freedom House	0.76
Global E-Government	0.53
Bertelsmann Transformation Index	0.56
Surveys of Individuals	
Latinobarometer	0.52
Afrobarometer	0.70

* Data refers to 2004

Kaufmann, Kraay, and Mastruzzi

Table 2: Non-Robustness of KS Halo Effects Regressions

*Table reports the coefficient on two-year average prior growth in a regression of the
indicated aggregate governance measure on prior growth, per capita income,
population, and schooling. The first panel replicates KS Table 1. The second panel
replaces per capita income with log per capita income, and a stock of human capital
measure with a flow measure of enrollment rates.*

| | *Replication of KS Specification* | | | | | |
	VA	PV	GE	RQ	RL	CC
1996 Coef	1.28	4.92	2.57	3.21	0.07	3.49
t-stat	0.89	2.81	2.30	2.27	0.06	2.54
1998 Coef	3.31	7.70	3.58	6.56	3.63	3.05
t-stat	1.78	3.98	2.53	4.01	3.05	2.47
2000 Coef	5.70	7.31	5.43	5.46	3.41	3.67
t-stat	2.74	3.42	3.42	2.98	2.59	2.33
2002 Coef	4.77	7.64	5.83	5.32	5.32	3.04
t-stat	2.25	2.86	3.42	3.26	3.17	1.73
2003 Coef	1.44	6.33	3.72	3.58	4.21	2.01
t-stat	0.66	2.77	2.34	2.18	2.68	1.27
2004 Coef	0.09	5.92	4.25	5.73	5.11	2.61
t-stat	0.04	2.83	2.82	3.73	3.45	1.70
2005 Coef	-0.23	5.24	4.25	5.77	4.62	2.83
t-stat	-0.10	2.47	2.83	3.76	3.10	1.86
# Significantly>0:	**32**					
# Significantly<0:	**0**					

| | *Results for 150 Country Sample* | | | | | |
	VA	PV	GE	RQ	RL	CC
1996 Coef	1.94	3.46	2.23	2.50	1.07	2.52
t-stat	1.91	3.17	2.48	2.86	1.24	2.46
1998 Coef	-1.41	5.03	0.05	-0.55	-1.25	-0.22
t-stat	-1.22	3.26	0.05	-0.52	-1.34	-0.22
2000 Coef	-1.25	2.87	-2.41	-1.43	-2.59	-2.85
t-stat	-1.18	1.79	-2.65	-1.38	-2.99	-2.93
2002 Coef	-1.01	3.84	0.69	1.34	0.12	-1.34
t-stat	-0.67	2.18	0.59	1.07	0.09	-0.97
2003 Coef	-3.09	2.52	0.39	0.76	-0.78	-2.44
t-stat	-1.93	1.57	0.32	0.59	-0.63	-1.85
2004 Coef	-3.49	0.43	-1.85	-0.19	-2.33	-4.72
t-stat	-2.19	0.31	-1.59	-0.15	-1.91	-3.67
2005 Coef	-3.72	0.49	-1.25	0.40	-2.37	-3.97
t-stat	-2.28	0.34	-1.06	0.33	-1.92	-3.03
# Significantly>0:	**6**					
# Significantly<0:	**7**					

Table 3: Non-Robustness of KS Halo Effects Regressions, Cont'd

This table replicates the regressions in KS Table 1 but adding prior 20-year average growth as a right-hand-side variable.

Dependent Variable	GE 1996	GE 1998	GE 2000	GE 2002	GE 2004	Evans-Rauch
Per Capita GDP	0.096 (0.010)***	0.087 (0.010)***	0.074 (0.009)***	0.068 (0.008)***	0.073 (0.007)***	0.240 (0.106)**
Human Capital	0.038 (0.031)	0.032 (0.034)	0.033 (0.031)	0.072 (0.028)**	0.052 (0.027)*	-0.190 (0.328)
2-year prior growth	0.689 (1.383)	1.935 (1.624)	3.234 (1.618)**	1.668 (1.946)	1.453 (1.643)	-6.100 (18.936)
20-year prior growth	5.851 (2.612)**	5.818 (3.008)*	9.934 (2.609)***	10.228 (2.655)***	9.236 (2.643)***	67.001 (27.238)**
log(population)	-0.009 (0.029)	-0.054 (0.032)*	-0.011 (0.030)	-0.033 (0.026)	-0.035 (0.026)	0.805 (0.385)**
Constant	-0.765 (0.485)	0.006 (0.540)	-0.692 (0.507)	-0.508 (0.441)	-0.440 (0.436)	-8.609 (6.916)
Observations	93	94	97	98	97	31
R-squared	0.86	0.82	0.83	0.87	0.87	0.53

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Kaufmann, Kraay, and Mastruzzi

Endnotes

---

[i] These governance indicators capture six dimensions of institutional quality: Voice and Acountability, Political Instability and Violence, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption. The indicators cover over 200 countries and are available for 1996, 1998, 2000, and annually since 2002. The latest update of the governance indicators is described in Kaufmann, Kraay, and Mastruzzi (2006b), and the data and a web-based graphical interface are available at www.worldbank.org/wbi/governance/govdata. Our earlier papers documenting these governance indicators are Kaufmann, Kraay and Zoido-Lobatón (1999a,b and 2001) and Kaufmann, Kraay, and Mastruzzi (2004, 2006a,b).

[ii] While not one of our main points, we find KS's discussion of the possibility that firms' perceptions of onerous and excessively bureaucratic regulation simply reflect their disgruntlement with a benevolent regulatory regime to be a bit naive. See for example Shleifer and Vishny (1999) for an extensive account of "grabbing hand" bureaucrats who use regulation to extract rents for themselves.

[iii] We use the individual or average of individual questions from each of listed sources as it enters into our aggregate governance indicators. Details on these specific measures can be found in Appendices A and B of Kaufmann, Kraay, and Mastruzzi (2006b).

[iv] See for example Glaeser et. al. (2004) who assert without evidence that the correlation between per capita incomes and perceptions-based measures of governance is due to such halo effects. In Kaufmann, Kraay, and Mastruzzi (2006a) we show that this argument is unlikely to matter much empirically because it would require implausibly large halo effects.

[v] Our results here are not exactly the same as those reported by KS because of minor discrepancies between our datasets. We use a slightly-revised newer version of our governance indicators, and also a slightly different measure of human capital from the Barro-Lee dataset. However, our results mirror exactly the pattern of significance that KS show and so we think constitute a "fair" replication of their finding.

[vi] Although this flow measure is of course conceptually different from the stock of human capital, the two are quite highly correlated across countries. And since KS do not offer a theoretical model which insists that the stock measure be used, we think it is reasonable also to look at the flow.

[vii] The variance share is defined as (VAR(halo effect) + COV(halo effect, government effectiveness))/VAR(government effectiveness). See Klenow and Rodriguez (1997) for a justification.

[viii] See for example Hall and Jones (1999) and Rodrik, Subramanian and Trebbi (2004) for other contributions to this literature.

[ix] This omission is particularly troubling given that they devote the entire middle part of their paper to arguing that there is in fact one such channel of reverse causation, through halo effects.

[x] Leading examples are the construction of creative instrumental variables, such as in Acemoglu, Johnson and Robinson (2001), the use of time-series exclusion restrictions such as in Chong and Calderon (2000), and the use of identification through heteroskedasticity as in Rigobon and Rodrik (2004).

[xi] KS ignore the difficulty of estimating growth regressions, which are intrinsically dynamic panel regressions, using ordinary least squares. As is well understood, in the presence of unobserved country effects, initial income is by construction correlated with the error term, and this endogeneity problem in general contaminates estimates of all of the coefficients of interest. See Caselli, Esquivel, and Lefort (1996) for an application to growth empirics. We also find KS's choice of estimation technique for their pooled regressions peculiar. In the last column of their Table 1 they choose a random effects estimator, while in the last column of their Tables 3 and 4 they choose a fixed-effects estimator. As noted above neither a fixed effects nor a random effects estimator will yield consistent estimates of the slope coefficients in the growth regressions in Tables 3 and 4. However, we find it interesting that had KS used the same random effects estimator in Table 3 as they had in Table 1, they would have found that government effectiveness *actually enters positively and significantly for growth in their own specification.*

Growth and Governance: A Rejoinder

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi
The World Bank[1]
December, 2006

(forthcoming, Journal of Politics)

In this brief rejoinder we would like to point several key areas where we disagree with Kurtz and Shrank's response entitled "Growth and Governance: A Defense".

Definitions

KS begin their defense by accusing us of "abandoning" our measure of Government Effectiveness in favour of other indicators.such as Rule of Law. All we noted in our original response, and reiterate here, is that we find KS's exclusive emphasis on this one particular dimension of governance to be idiosyncratic, and not shared by the large economics literature on institutions and growth. We also note that we discussed all six of our governance indicators in our response simply because the critiques that KS raise of government effectiveness, notably potential respondent biases and halo effects, could equally well be applied to our other indicators. We think it important to demonstrate the scarcity of empirical support for these critiques for all six of our governance indicators.

Ideological Biases

---

[1] 1818 H Street NW, Washington DC 20433, dkaufmann@worldbank.org, akraay@worldbank.org, mmastruzzi@worldbank.org. The opinions expressed here are the authors' and do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent.

Kaufmann, Kraay, and Mastruzzi

KS remain unpersuaded by our arguments that ideological and business-oriented biases of respondents do not play a significant role in our indicators. Regrettably, however, KS continue to fail to provide any concrete evidence of such biases. KS quote the opinions of some scholars that some of our data sources are "ideologically loaded". As we noted in our response, in the case of ideological biases, we have empirically investigated whether supposedly "right-wing" think tanks give better scores of governments who share their political orientation, and found virtually no evidence of this. Any empirical evidence to the contrary that KS could bring would be a welcome addition to the debate, but simply citing the assertions of others strikes us as unhelpful.

The evidence KS report from the Afrobarometer surveys on businesspeoples' preferences is irrelevant to their critique of our indicators. We acknowledge that KS have a useful suggestion that one can use the identity of respondents to this survey (who are asked to self-identify as businesspeople or other types of respondents) to investigate whether businesspeople's views differ markedly from others' on what constitutes good governance. All KS have shown is that on average businesspeople respond more favourably to questions about government services. While this is interesting, it is irrelevant to the issue at hand. What matters for our indicators is how countries are ranked relative to each other. Therefore, KS would need to show that a ranking of countries based on businesspeoples' responses differs significantly from a ranking of countries based on other responses. It is entirely possible that businesspeople on average in all surveys provide more positive responses than other respondents (perhaps they are

Kaufmann, Kraay, and Mastruzzi

on average just more optimistic?). But this need not affect the relative ranking of countries, which is what matters for our indicators.

KS also emphasize the relatively low correlation of the Latinobarometro survey question we use in the Government Effectiveness Indicator with corresponding variables from commercial risk rating agencies. It is useful however to recognize that the particular Latinobarometro question we use for Government Effectiveness is an unfortunately vague question about respondents' overall trust in government. In contrast, the Afrobarometer survey asks several questions much more specifically focused on access to public services that are much closer to those asked of firms in the Global Competitiveness Report survey, and these two sources accordingly are much more highly correlated. Indeed, if we compare a more specific question about trust in police from Latinobarometro, that we use in our Rule of Law indicator, with a similar question about the police from the Global Competitiveness Survey, we find a quite high correlation of 0.77. We are therefore neither surprised nor particularly concerned at the lower correlation of the one particular -- and unfortunately vague -- Latinobarometro question that KS emphasize in their Defense.

We do not dispute that data sources capturing the views of businesspeople and commercial risk rating agencies play a prominent role in our governance indicators (although certainly not an exclusive role, given our reliance on citizen surveys, NGOs, and multilateral organizations). We note also that this feature of our indicators is dictated by data availability -- there simply are not many regularly-updated cross-country

household surveys that we can incorporate in our indicators. However, the key issue is whether country rankings based on such alternative and yet-to-be-created data sources would be substantially different from those we present. On this key point KS provide no new evidence.

Halo Effects

We can only read KS's Defense on this issue as a conceding our point that their original work failed to provide robust empirical evidence of halo effects. KS nevertheless wonder why we "depart from past practice" by using secondary school enrollment rates to expand the sample of countries (selectively citing one or two of our earlier papers on completely different topics), and suggest that the lack of robustness of their results in fact constitutes a critique of our indicators. In response, we briefly note that:

- Although of course conceptually distinct, secondary schooling rates are extremely highly correlated with stocks of human capital, with a correlation in our sample of 0.85. While we are unaware of any well-specified theoretical reason to prefer the stock variable over the flow variable as a determinant of government effectiveness, it is perfectly reasonable to use the flow data as a proxy for the stock in the interests of expanding country coverage. And as we showed, in this larger sample there is no systematic evidence of halo effects.

- As KS themselves note, enrollment rates could well be interpreted as an indicator of government effectiveness itself. This is a useful observation that in fact provides a further justification for using this variable, since it serves as an "objective" control for unobserved true government effectiveness. But under this interpretation, KS wonder why enrollment rates are not more strongly correlated with government effectiveness. The answer is straightforward, as the regressions KS propose also include per capita income together with schooling variables as explanators of government effectiveness. Unsurprisingly these two right-hand-side variables are very highly correlated at around 0.8, and this multicollinearity problem makes it difficult to identify a significant *partial* correlation between schooling and government effectiveness. In fact, the simple correlation between the two is a very respectable 0.72.

Growth and Governance

KS assert that our measure of Government Effectiveness should by construction be positively correlated with short-run subsequent growth, claiming that our measure contains several questions about anticipated short-run growth. This claim is based on an overly-literal reading of the criteria used by just one of our 17 data sources for Government Effectiveness. DRI, a commercial risk rating agency, uses a somewhat peculiar phrasing when it provides numbers which show their assessment of the likelihood of certain "risk events" like "a decrease in government personnel quality that lowers growth in a 12-month period." We have discussed this question at length with

staff of DRI and have come to the understanding that this is simply peculiar phrasing serving their own marketing purposes, and that they are actually providing an assessment of levels of civil service quality. This, combined with the sensible observation that we do not think commercial risk rating agencies have perfect foresight about growth even one year hence, means that we are completely unconcerned that this particular measure, or the aggregate indicator, are uncorrelated with subsequent very short-run growth in the handful of empirical specifications that KS offered in their article.

We also think it important to correct the mischaracterization of the empirical literature on institutions and growth that KS offer in their Defense. In particular:

- KS claim that the idea that differences in per capita income today primarily reflect differences in long-run growth performance requires "heroic assumptions about the distribution of income across countries in 1800". This stylized fact is not an assumption, but rather is based on the careful studies by economic historian Angus Maddison, who documents as carefully as possible that the ratio of per capita incomes between the richest and poorest countries in the world 200 years ago was on the order of three or four to one, while today it is an order of magnitude greater. Based on this factual observation it logically follows that the much larger cross-country income differences we observe today primarily reflect very long-run differences in growth performance.

- KS claim that using perceptions of expropriation risk is using a measure of policy (the decision to expropriate) instead of an institution. This is a rather superficial

interpretation, and would only be true if studies used actual expropriations, rather than assessments of the likelihood of expropriations averaged over a decade. Investor perceptions of expropriation risk are widely interpreted as capturing something about the respect of the state for private property, which is a fundamental notion of institutional quality.

- KS dispute the use of settler mortality rates as an instrument for institutional quality because settlers were "the most violent expropriators in world history". It appears to us that KS are choosing to miss the point of one of the most influential empirical papers on institutions and growth. That Europeans propagated all sorts of atrocities in the countries where they settled is neither in dispute, nor is it in any way relevant to the empirical strategy at hand. What matters is that the disease environment faced by settlers provides an exogenous source of variation in the incentives that settlers faced to set up the institutions of property rights protection, which persist to this day. This exogenous variation can be used to statistically identify the causal effect of institutions on very long run growth.

As we noted in our original response, we do not wish to argue that the empirical literature on institutions and growth is conclusive -- and so we cannot agree with KS's characterization of us as taking an "intolerably conservative approach to social scientific argumentation". We think it is entirely appropriate for KS, or any other scholars, to critique, and so develop, the literature on institutions and growth. But we do not think that offering a few regressions, so far removed from the best-practice frontier in this

literature, constitutes a serious critique that helps to advance our understanding of these important questions.