Governance Indicators
Some Proposals*

Piero Stanig [†]

Mark A. Kayser [‡]

2013

Abstract

Good governance extends beyond narrow questions of efficiency and of control of corruption, restrictively defined as the use of public office for private gain. Nevertheless, many governance indicators focus on narrow issues, in particular individual malfeasance, to the exclusion of broader questions about effective governance understood as the precondition for the delivery of the most welfare to stakeholders at the least cost. Broader measures of governance, on the other hand, often conflate governance with outcomes such as national income — a problematic construction given that researchers often wish to investigate the effect of governance on economic outcomes. We suggest that understanding the how (or how well) of governance requires also understanding questions about in whose interest a government governs. This chapter broadens the definition of misgovernance to include those often fully legal activities, structures and decisions that reduce the utility of a broader set of legitimate stakeholders to the benefit of smaller sets of individuals and organizations. We propose a set of novel governance indicators that operationalizes two missing aspects of extant measures: organizational capture and the capacity for governance innovation.

Chapter 9

Governance Indicators: Some Proposals

Piero Stanig and Mark Kayser

> *'If you cannot measure it, you cannot improve it.'*
>
> *Lord Kelvin*

Governance is central to human welfare and, as such, it has been subject to intellectual enquiry since antiquity. Yet only since the 1990s, partly due to the institutional turn in economics, have sophisticated cross-national measures of the quality of governance emerged. What the fledgling field measuring the quality of governance lacks in age, it has made up for with quantity: in less than two decades, numerous measures of the quality of governance, covering aspects from corruption to business regulation to respect for human rights, have materialised. We argue here that the field is in need of a systematic assessment and that future measures of governance—such as those that we propose below—can contribute much to the study and practice of governance by combining a broad theoretical approach with methodological innovation.

A naïve observer of the abundance of governance indicators might reasonably conclude that researchers and policymakers command the measures necessary for designing and evaluating efficient and effective governance. Closer inspection of the literature, however, reveals overrepresentation of specific aspects of governance and types of measurement and the systematic neglect of others. Many of the available governance indicators focus on narrow issues, in particular the control of corruption, restrictively defined as the use of public office for private gain, to the exclusion of broader questions about effective governance that enables the delivery of the most welfare to stakeholders at the least cost, but more importantly about the capacity to adapt to new challenges.

The narrower measures might miss many of the ways in which inefficiency and ineffectiveness in governance arise. For instance, poor governance might be a consequence of deliberate mis-design of policy in contexts in which

189

Piero Stanig and Mark Kayser

powerful special interests 'capture' the state. Alternatively, they might be driven by incompetence and a lack of the intellectual resources—within and outside of government apparati—needed for innovative policymaking and, when necessary, reform. Governance failures might also stem from the absence of a sound civil society that contributes to the identification of problems, monitors policy implementation, and fosters innovation. Finally, as suggested in Anheier et al. (2013), governance failures, in the contemporary globalised setting, might follow from the reluctance of nation states to cede some sovereignty and cooperate in the production of global public goods.

While these phenomena are not addressed directly by the narrower measures of governance, many extant measures that avoid the pitfall of narrowness are closer to measures of prosperity than measures of governance proper (e.g. the Ibrahim Index of African Governance, IIAG), and some, while in essence measures of governance, also throw into the mix some elements of prosperity (e.g. the World Economic Forum's Global Competitiveness Index and the Bertelsmann Transformation Index, BTI). The inclusion of outcomes that are plausibly more usefully thought of as consequences of governance, rather than part of governance's conceptual definition, renders such broader measures of governance problematic when it comes to testing claims derived from theory—for instance, about the consequences of governance on prosperity, development, or equality.

In this chapter, we briefly review the state of the governance indicators field: what is measured, what is mismeasured, and what is not measured. Key to this exercise is the observation that, regardless of their intention, many more measures of *government* than of *governance* populate the field. We argue that effective and efficient *governance* depends not only on sets of institutions that constitute measures of *government*, but also on the balance of influence between interests and those institutions. Two aspects of this definition of governance are (a) the independence of policymakers from the interests they regulate and (b) the analytic and innovative capacity that governments can draw on to identify and construct solutions to problems. Both connect government to society, but neither is systematically measured. Our purpose in this chapter is to propose and motivate a set of indicators that will better capture the concept of effective and efficient governance. As such, this chapter provides no exhaustive survey of the field—there are several others that do that.[1] What we do provide is a substantive and methodological road map to developing a dashboard of *governance*.

---

[1] Broad surveys of governance indicators are provided, among others, by UNDP (2004) and Bandura (forthcoming). Online and more readily updated surveys are available on the websites of the Quality of Government Institute <http://www.qog.pol.gu.se/> and the American Political Science Association Democracy Audits and Governmental Indicators project <https://sites.google.com/site/democracyaudit/>.

Governance measures: A Brief History

Attempts to quantify governance quality have grown quickly of late. Until relatively recently, there were few attempts to measure the quality of governance quantitatively and phenomena related to it like bureaucratic efficiency and corruption. Several broad measures in other social domains preceded the emergence of dedicated governance indicators. Measures of prosperity, such as the United Nation's Human Development Index, have been highly prominent, and the field is developing at a rapid pace. Proposals to replace narrow measures of prosperity such as gross domestic product (GDP) with broader alternatives (see, for example, Stiglitz et al. 2010) echo, to an extent, those made over the decades by the 'social indicators' field (Hagerty and Land 2007) and incorporated, for instance, in the Canadian Index of Wellbeing.

It was only with the interest of private rating agencies that sell their evaluations to potential investors that dedicated measures of governance—or aspects of governance—started to emerge in the 1980s. Corruption featured prominently among these. The turning point in the quantification of governance came when scholars began to employ these ratings in academic research, in particular development economists who wanted to assess the role of governance in development. This reflected the new interest displayed by political economists in the role of institutions, and the interest, in economics and in political science, in the systematic study of corruption. Before that, economists tended to relegate political institutions to a minor role, and research on corruption in political science took more of an ethnographic bent. In seminal papers, Mauro (1995) first used the indices of institutional quality provided by the Economist Intelligence Unit (EIU) to predict economic growth, and Knack and Keefer (1995) used measures from the International Country Risk Guide (ICRG) and the Business Environment Risk Intelligence (BERI) reports to predict growth and investment.

At that point, the need for some rigorous measures of governance became apparent. In 1996, the World Bank Institute launched the Worldwide Governance Indicators (WGI) that rate countries on six dimensions. The WGI are based on the aggregation of several individual indicators, some of them based on subjective expert judgements, some on survey data. The main reason for their prominence is twofold. First is their reputation for rigour which rests on the academic credibility lent by the direct involvement of professional economists, the transparency of the methodology adopted to estimate both the scores and their uncertainty, and the compelling presentation of the indicators themselves. Second, their perceived affiliation with the

191

Piero Stanig and Mark Kayser

World Bank gives the indicators both an aura of officialdom and a taint of 'globalisation as imperialism'.[2]

A substantively and methodologically different approach emerged in 2003 with the World Bank's Doing Business (DB) indicators.[3] Based on a method first developed in Djankov et al. (2002), these measures employed a form of field experiment to estimate barriers to entry in ten areas across 183 countries. Rather than consulting experts for subjective estimates of broad and difficult to measure concepts like corruption, the Doing Business project objectively measures the time, cost, and difficulty of several procedures associated with business regulation and enforcement. By comparing the difficulty of complying with and completing tasks—for example, incorporating a new business—across countries, the project compiles a single international (and more recently, subnational) index of the 'ease of doing business'. The advantage of this approach lies in its internal and (quite narrow) external validity. The drawback, however, is that extrapolations from these narrow measures to even slightly broader and related topics, such as 'quality of regulation', jeopardise their original validity.

Despite its influence, DB's methods remain more of an exception than the rule. Like Doing Business, many current governance indicators incorporate objective components—albeit not via field experiments. But where DB's objective components relate directly to governance, most others' relate to an associated construct: prosperity. Prominent examples of such measures include the World Economic Forum (WEF) Competitiveness Index, the Bertelsmann Transformation Index, the Legatum Prosperity Index, and the Ibrahim Index of African Governance. We call these indicators 'hybrid' because they combine assessments of institutional quality with measures of prosperity (or development) *outcomes*.

The Quality of Government (QoG) dataset (Teorell et al. 2011), which is without doubt the most comprehensive dataset with country-level data related to governance, usefully divides variables into three categories: 'What It Is' variables pertain to governance proper; 'How To Get It' variables, potential causal antecedents of good governance; and 'What You Get' variables measuring prosperity outcomes. What we call 'hybrid' measures conflate 'What It Is' with 'What You Get' variables. In the Ibrahim Index of African Governance, for instance, indicators of rule of law, corruption, and fairness of elections (all measures of institutional quality) are averaged with measures

---

[2] A similar effort came also from a civil society organisation, Transparency International, which since the late 1990s has been releasing, first sporadically, then annually, the Corruption Perceptions Index. This is very similar to the WGI 'graft' (later 'control of corruption') dimension, in terms of data sources, *which rely mostly on expert assessment scores*, but simpler in terms of methodology.

[3] <http://www.doingbusiness.org>

of health and education outcomes, soundness of macroeconomic policy (e.g. ratio of debt to revenue), and infrastructure provision, which are clearly development outcomes. Similarly, the WEF Competitiveness Index aggregates, among many others, some measures of institutional quality and market efficiency (which can be considered causes of prosperity, and possibly components of governance) with measures of health and education, which are outcomes of good governance and better understood as components of prosperity.

In a similar fashion, the Bertelsmann Transformation Index measures the progress of countries towards the goal of resembling Western European countries in political and economic terms. In the words of the organisation, 'the BTI focuses on the normative goals of democracy under the rule of law and a market economy anchored in principles of social justice in its analysis of political decision makers targeting these goals...'[4] Bertelsmann produces three different indices. The first two measure 'political transformation' and 'economic transformation'. The third, the 'management index', introduced later, focuses more on the 'quality of governance' at a relatively high level of abstraction. Political transformation is measured based on criteria like 'stateness', political participation, rule of law, stability of democratic institutions, and political and social integration. Economic transformation is measured based on criteria among which are the level of socio-economic development, the organisation of the market and competition, currency and price stability, economic performance, and sustainability. The management index deals with more abstract criteria, such as steering capacity, resource efficiency, consensus-building, and international cooperation. The questions are answered by experts drawn from leading universities and research institutions around the world. An approach analogous to the BTI, based on a network of experts and centralised review of the evaluations, has been recently adopted by the Global Integrity Report, which focuses on issues such as transparency of the public procurement process, media freedom, asset disclosure requirements, and conflicts of interest regulations.[5]

The available indices focus, in a sense, much more on government than governance. In our perspective, governance involves multiple social actors, and good governance requires a balance of power (or the 'right mix') between state and non-state actors. The issue of the state–society relationship can be approached from different perspectives. On the one hand, a government

---

[4] <http://www.bti-project.org/index/methodology/>

[5] A notable example of a measure of governance that, at the same time, employs objective measures but avoids too narrow of a definition of governance, is the Revenue Watch Index. It provides an 'assessment and comparison of information published by governments about revenues, contract terms, and other key data' based on an aggregation of publicly available information about natural resource management.

Piero Stanig and Mark Kayser

that is too weak vis-à-vis private interests might become 'captured' and pursue particularistic objectives, rather than focus on public good provision or equalisation of opportunities, provision of safety nets, or similar imperatives. But a government that is too insulated from society, and specifically from knowledge producers (e.g. academia and think tanks) and from advocacy (civil society), might lack the intellectual resources to solve new challenges and might not be able to collect information about these new challenges (to the point of even ignoring their existence). In the sections below devoted to our 'effectiveness and efficacy' dashboard, and in our discussion of state capture, we explain in more detail the implications for empirical research of this perspective.

Composite Indices versus Dashboards: Some Methodological Considerations

A fundamental question when developing indicators is how to present them. Broader constructs are often only partly captured by a single measure, leading to plural measures for many phenomena of interest. Scholars often address this problem by building composite indicators. For example, the United Nation's Human Development Index aggregates individual indicators measuring life expectancy, education, and income. In this chapter we advocate for the presentation of indicators in the form of 'dashboards'—i.e. sets of key indicators related to a broader concept of interest. The communicative advantage of dashboards relative to composite indices is clearly conveyed by the well-known analogy used by Stiglitz et al. (2010): If you want to know how well your car is performing, it makes sense to read the fuel gauge as well as the speedometer. No single indicator—and also no composite index—will provide the information one needs. Methodologically, dashboards also obviate the need for certain trade-offs related to the construction of composite indices. In this section, we explain why we favour this approach, and to do so we provide a brief methodological overview of governance indices and, in general, of estimation of composite indices.

Aggregation

Most of the efforts to provide quantitative measures of governance estimate a composite index arrived at through the aggregation of several different pieces of information. Aggregation serves different purposes, and, in particular, three broad functions of aggregation can be isolated:

194

1. *Signal extraction*: from several different imperfect or indirect measures of a given concept, an aggregate index 'separates the wheat from the chaff'.

2. *Noise reduction*: in general, the aggregation of repeated independent measurements of the same quantity leads to a measure that is more precise (less variable) than each of the individual measurements.

3. *Scaling*: different measurements that 'live' in different metrics can be reduced to a common metric, making broader comparisons possible.

It is worth starting the discussion with an overview of the most sophisticated class of methods to construct aggregate indices and a discussion of methods that, at face value, seem 'simpler' in the context set up by the complicated models. In other words, it is worth looking at the most complicated way of aggregating in order to throw light on what the functions of aggregation themselves might be. For the sake of clarity, we use the name 'factor analysis' to refer to the sophisticated methods, without specifying the subclass to which they belong. In the field of governance measurement, the most prominent effort that relies on this approach is the World Bank Institute's WGI.

Crucially, all the methods that can be classified as 'factor analysis' perform, substantially in one single step, the basic functions listed above: they extract a signal from noisy indicators; they provide ways to aggregate, hence reducing the variability of the index; and finally, they scale different sources so they can be meaningfully combined. These different functions are performed pretty much 'in one shot' by factor analysis models. The model is estimated to extract underlying ('latent') dimensions based on the covariance of observable variables, hence allocating part of the variation in the observable data to signal (the latent factors) and part to noise. The estimation also provides weights ('factor scores' or 'discrimination parameters') that can be used for aggregation (via weighted averaging) of different observable variables, hence providing guidelines to increase precision and to scale different observables on the same metric.

Factor analytical models, in their general form, rely on a seemingly technical assumption—the assumption of 'local independence'. The idea is that, conditional on the true (and unobserved) value of the latent phenomenon, the errors in the various indicators are independent. Seen from one perspective, this assumption is the one, traditionally stated, that offers the reason why different variables, which measure different aspects of a phenomenon, covary: their dependence on the latent phenomenon. More weight is given to those variables that covary more with each other, precisely because it is from such covariance that the information about the latent phenomenon comes. But the assumption of local independence can also be seen, as an intuition, from a different angle: namely, the assumption implies that *the*

195

Piero Stanig and Mark Kayser

only reason why two observable variables covary is the fact that they both depend on the latent phenomenon (and therefore they are assumed to not covary for any other reason). This assumption is obviously questionable: both its realism and the practical implications of its lack of realism depend on the specific application. We discuss the two issues (realism and practical implications) in turn.

The realism (or empirical plausibility) of the assumption depends on the type of latent concept one is interested in measuring, and on the type of observable indicators one can collect to attempt to measure the concept. One can easily conjecture that in the realm of measurement of unobservable but theoretically fruitful political–economic concepts like 'governance' via aggregation of several observable political–economic phenomena, the assumption of local independence is much less tenable than it might be in other fields—like, for instance, psychometrics (Bollen 2002), social psychology (e.g. Schwartz 1992), and psychology of personality (e.g. Digman 1990). For instance, it is not far-fetched to claim that scores on different intellectual ability tests covary within individuals because they are the consequence of some latent phenomenon, called intelligence. For that matter, one can define intelligence as that ability that allows a given individual to correctly solve a diverse array of mathematical and verbal puzzles. In the world of political economy, to which the governance concept belongs, this is much more problematic.

This point has been forcefully made by Breusch (2005) in his criticism of so-called multiple indicators multiple causes (MIMIC) models for the estimation of the size of the 'shadow economy' (Giles and Tedds 2002; Bajada and Schneider 2005; Dell'Anno and Schneider 2006). Like some aspects of governance (e.g. corruption), the size of the 'shadow economy' is, on the one hand, linked tightly to many (if not all) important political–economic phenomena, and on the other, very hard to observe directly. The models used to estimate the shadow economy are, basically, variants of factor analytical models with a structural–equation modelling aspect: they try to back out the size of a latent variable (the shadow economy) by exploiting observable variables, some of which are causally antecedent, some causally consequent, to the latent phenomenon. Compare this with IQ testing, in which the latent phenomenon, intelligence, is assumed to be causally antecedent to all observables. An analogy to the MIMIC model in psychometrics would be if IQ testing involved neurological or anthropological measures—say, electrical activity of the brain and cranial volume—as causal antecedents of IQ itself.

It is worth noticing one fact that is usually overlooked. As long as the assumption of local independence (or conditional independence of the errors) is tenable, a factor–analytic approach should work better—in the sense of leading to results that might be more believable—if the observed indicators

196

are substantively very disparate. If one is able to detect correlations between different variables, which cannot be related to each other for any reason other than the underlying latent phenomenon one is interested in measuring, one can have more confidence in the resulting findings about the latent phenomenon. But for this to be true, something similar to a set of 'exclusion restrictions'—i.e. assumptions about which causal relationships between variables *do not* exist—is required. It has to be plausible to assume that observable indicators Y_1 and Y_2 cannot cause each other, either directly or indirectly, and that they cannot be consequences of a common cause other than the latent phenomenon one is aiming at. Yet, in the context of a political economy concept like governance, serious worries have to be raised about the kinds of exclusion restrictions that can plausibly be made. The application of factor analytical models to political–economic measurement, without paying attention to the potential relationship between political–economic observables, and therefore to the implicit 'exclusion restrictions', can be problematic. In psychology, the assumption of local independence might be often tenable, but in an economic setting it is much less tenable.

Returning to Breusch (2005: 28), in MIMIC models of the 'shadow economy', conditional independence means substantively that observable indicators like GDP and currency holdings are assumed to be related to other observables, like tax rates and unemployment rates, only through the size of the underground economy. Once the substance of the assumption is made explicit, its lack of realism is apparent. Crucially, unrealistic assumptions are not in themselves troubling, as long as they are innocuous. But in the case of a composite index, overlooking the violations of the 'exclusion restrictions' has a serious consequence. Specifically, the practical implication of implausible local independence assumptions (with their implicit exclusion restriction) is to attribute to the latent phenomenon (be it the 'shadow economy', 'governance' or 'intelligence') also the role played by all the other direct and indirect causal connections between the variables in a given set of observable indicators.

This concern is particularly pressing if the latent concept is narrowly defined and the observable indicators are broadly defined. One can well assume that, for instance, opinions about corruption among foreign investors and opinions about corruption among the local public are correlated with each other because they both depend on the underlying phenomenon of corruption: the latent phenomenon ('corruption') is relatively broad, and the observables are relatively narrowly defined. The assumption is less tenable were one to claim, for instance, that exports and foreign direct investment (FDI) inflows are related to each other only because of 'bureaucratic red tape'. It is not difficult to think of ways in which exports and incoming FDI are related to each other that have nothing to do with 'red tape'. Were one

197

Piero Stanig and Mark Kayser

to go ahead and make the assumption required to estimate a latent variable model, one would attribute to 'red tape' the consequences of several other phenomena that jointly affect (or are affected by) both exports and FDI.

A direct example of the problems with exclusion restrictions is provided by one of the criticisms often waged against governance indicators: that they suffer from a 'halo effect'. Countries that experience a period of good economic performance (for reasons unrelated to governance proper) might be perceived, by the public and by experts, as being better governed than before the economic boom.

To sum up: in the complex web of causal connections that link political–economic phenomena, it might be very difficult to isolate empirical measures that covary *only* because they are related to governance. But doing so is a prerequisite for the extraction of a latent measure of the quality of governance. Disaggregated measurements in a dashboard that throw light at a phenomenon from different perspectives avoid this problem.

Weighting

Another function performed by factor analytical models is to 'select' the relative weights that should be assigned to different observable phenomena when computing a composite index. Assigning weights is the same as making claims about the relative importance of each variable: in some way, to define what the concept is.

For instance, assume one were to create an index of 'regulation of economic activity' by collecting information about several ways in which governments restrict the choices of economic agents. Among these, one might consider also including, for instance, measures of the strictness of anti-discrimination laws that prevent employers from deciding whether to hire or fire an employee based on such criteria as somatic traits or religious affiliation. If one were to assign equal weight to this variable and to an indicator of prohibitions on polluting rivers, the underlying concept captured by the indicator would be radically different from the case in which a very small weight (or none at all) were assigned to anti-discrimination strictness, and similarly, radically different from the concept implicitly defined by an index that assigned no weight to environmental protection regulation and all the weight to anti-discrimination laws. Hence, a seemingly technical decision about weighting has very important consequences for what is meant by 'regulation'.

From a more analytical point of view, the choice of weights is equivalent to making a set of claims regarding the slopes of a set of trade-offs. By saying that 'governance' is estimated, for instance, by the sum of (a score of) 'control of corruption' and (a score of) 'rule of law', we are implicitly claiming that control of corruption and rule of law are substitutes, and that there exists an

198

'iso-governance curve', or a curve, in the two-dimensional space defined by control of corruption and rule of law, such that governance remains at the same level if control of corruption is decreased by some amount and rule of law is increased by some other amount. The relative amount of rule of law required to compensate a decrease in control of corruption is specified by the weights assigned to control of corruption and rule of law. The weights, then, are much more loaded with substance, in terms of claims about the functioning of the real world, than one might suspect. Factor analytical models provide a data-driven method of choosing relative weights. Outside of factor analysis, though, the choice of weights has to be based either on extra-data knowledge, or on data that do not enter into the estimation of the composite index directly, which usually means information on a phenomenon of interest (e.g. economic performance) that the latent phenomenon (e.g. governance) should be able to predict. We discuss the latter approach first.

One strategy to which some index projects have resorted in order to decide on which weights to assign to different sources is that of predictive power for some outcome of interest. The outcome of interest is regressed on the available observable indicators, and the partial regression coefficients from such a regression are treated as weights. For instance, the WEF Global Competitiveness Index weights different variables according to the coefficient estimates in a regression of GDP per capita on the 'sub-indices' (partially aggregated subsets of observable indicators), and the Legatum Prosperity Index assigns weights based on the coefficient estimates of regressions with income and well-being as dependent variables. Clearly, this strategy, while not necessarily problematic at a general level, leads to the creation of indices that have very little value when used to test theory-driven hypotheses. It would not, for example, be surprising if the WEF Competitiveness score were to predict cross-national patterns of economic performance well. This would not, however, support a testable hypothesis like 'competitiveness is good for the economy'. The measure of competitiveness itself is designed to be predictive of economic performance by assigning more weight, in its estimation, exactly to those variables among the many collected that are more predictive of economic performance.

This criticism of the regression-based approach to weight choice works along the same lines as Kaufmann et al.'s (2010) rebuttal of Thomas (2010), who criticises the WGI because, in her view, it lacks 'convergent validity'—in other words, it is not constructed with the predictive power for other phenomena in mind. The criterion of convergent validity was originally proposed in psychometrics. As Kaufmann et al. (2011) explain, if the primary purpose of data collection is to assess the empirical relevance of theories (as it often is in political economy), then choosing measures that are consistent with prior theorising would lead to confirmatory bias. They continue by

199

Piero Stanig and Mark Kayser

stating that the lack of popularity of the 'machinery of construct validity' in economics is a consequence of the importance assigned to data analysis in discriminating among competing hypotheses.

Another approach one can follow to choose weights is assigning equal weight to each of the observable indicators that are included in the estimation of the index. As noted above, choosing weights means deciding about the slope of a trade-off (or 'iso-governance curve'): equal weights are not agnostic about the trade-off; instead, they assume unit-per-unit substitution between the different phenomena whose measurements are included in a composite index. The regression-coefficient weights do have the advantage of empirically estimating the trade-off by assuming a linear relationship between the concept that is being measured (in the case of WEF, competitiveness) and economic performance. It becomes apparent that equal weighting of different variables, far from being equivalent to choosing 'not to weight' or to 'not choosing', represents a very specific choice of weights, in which one assumes that one unit of control of corruption is the substitute for exactly one unit of rule of law.

One further point, closely related to the one made in the previous paragraph, and similar to a point forcefully made by Stiglitz et al. (2010), is that a single dimension onto which the different measures map might not always exist. In the automobile dashboard metaphor in the Stiglitz et al. (2009) report, and the subsequent Stiglitz et al. (2010) book, the authors note that an indicator that were to report the sum of the speed and the fuel level would be completely useless in the decision making of a driver because the two do not map onto one single dimension. Similarly, creating composite indices of governance from variables for which a trade-off makes very little sense, and therefore no reasonable mapping onto a single dimension exists, might be a pointless exercise. How much torture and how many extrajudicial killings are needed to compensate a unit decrease in the illiteracy rate? Would a measure of 'governance' that averages measures of the two phenomena meaningfully capture something that can be called 'governance'?

We present this discussion about weights for one specific reason: it directly points to an advantage of a 'dashboard' approach to governance measurement—that is, choices about trade-offs in this case are not made by the compiler of the dashboard. Instead, the available information is presented in disaggregated form, and users of the dashboard can use their knowledge of (or their conjectures about) the trade-offs as needed.

The Case for Dashboards

'Simple' methods of aggregation, far from being assumption-free as one might hope, are only superficially simple and demand a strong belief in

perfect unit-by-unit substitution between different observable indicators or components of the governance index in question. Aggregation might be meaningless when iso-governance curves are not well-defined due to multi-dimensionality of the theoretical constructs themselves. Factor-analytical models, while very powerful, have to be carefully constructed to avoid the pitfall of ignoring equilibrium effects and, in general, of the fact that in the context of political economy concepts, almost every phenomenon of some importance is, in one way or the other (e.g. causally, directly, or indirectly), related to other phenomena. Furthermore, weighting by predictive power for a given outcome might lead to creating measures of governance that have little or no value when attempting to analyse empirically the role of 'governance' for outcomes of interest and testing propositions about the causal role of governance in development, responses to external shocks, or other outcomes.

For all these reasons, and even if we are not as sceptical of aggregation as we might seem from the discussion above, we think there is value in also presenting full dashboards, along with summaries based on aggregation. The estimates of individual aspects in the dashboards can be used to assess, at a glance, the comparative performance of different countries in each aspect, and can also be used as inputs, if needed, in the construction of indices that are more aggregated than the ones we present.

As discussed above, there are different purposes of a governance measure, each permitting and possibly requiring different combinations of components in a composite measure. For example, one might want to create a measure of governance in order to predict future development outcomes. Alternatively, one might want to evaluate policy reforms (introduced in some countries and not in others) or some (plausibly exogenous) change in one observable phenomenon, and to assess their effect on governance. Or one might want to detect interesting correlations between good governance and desirable outcomes (like economic growth, income equality, or public health). More ambitiously, one might aim at estimating the causal effects of good governance on those same outcomes. Each one of these purposes requires, to an extent, a different definition, and a different operationalisation, of governance. In particular, a measure built following statistical learning principles (Hastie et al. 2009) to maximise predictive performance (for instance, by optimising predictive performance on data split between a 'training set' and a 'test set') might be very effective for forecasting purposes, but is useless if one wants to test hypotheses derived from theory (for example, about the effect of governance on economic development).

The dashboard approach provides information that can be used to create aggregate indices that fit the purpose of the specific analysis one is interested in. Moreover, it avoids the issue of conflating, on a single dimension,

Piero Stanig and Mark Kayser

inherently multi-dimensional phenomena, assuming the existence of an 'iso-governance curve' when this is, instead, not well defined.

At the same time, the dashboard approach has limitations that we fully appreciate. In particular, presenting a dashboard—rather than making the hard choices required to estimate a single index—might seem like an abdication of the analyst's responsibility to defend and justify one's conceptual and operational decisions. In the extreme, collecting dashboards might lead to the assembly of useless sets of heterogeneous measures without providing any rationale for the choice of inclusion of a given item in the dashboard. We are aware of these issues and they govern our drive towards reasonable parsimony.

We also propose—and in *The Governance Report* that is a companion to this volume (Anheier et al. 2013), provide—aggregate indices of medium-range concepts, based on what we believe, and explain to be, plausible aggregation schemes. In the next two sections, we describe proposals addressing two specific topics that draw our interest due to their neglect in the field of governance indicators. Both emphasise government's connection to the interests found among the governed: (1) the 'capture' of state bodies by private interests, and (2) the capacity of government, with the aid of external bodies, to identify, design and implement innovative policy responses.

Proposal 1: Measuring State Capture

We suggest that understanding governance requires understanding in whose interest an organisation—be it a government, a corporation, or a non-profit organisation—is governed. Misgovernance includes those—often fully legal—activities, structures and decisions that reduce the utility of a broader set of legitimate stakeholders to the benefit of smaller sets of individuals and organisations. Through forms of organisational 'capture', special interest groups exploit their privileged position to influence decisions and to bias outcomes to their advantage. This lowers the quality of governance by precluding policies that might enhance the overall effectiveness and efficiency of governance, but contravene the interests of privileged groups. Policy-making organisations that are captured by particular interests are more resistant to change, more inured to innovation, and less able to respond to challenges.

Here we explore two ways—a direct and a 'second-order' path—through which capture by interests in privileged positions affects governance. The direct path is simply that the interests catered to are narrower than those that, in an ideal state, would benefit from neutrally designed policy. The indirect, or 'second-order', path is a consequence of the concern among

202

those occupying privileged positions of a redistribution of power following effective governance: the sub-optimal, but certain, present arrangement is preferred to an improved, but uncertain, future innovative arrangement. Uncertainty about who is going to win and lose from reform can lead to a status quo bias, even with an egalitarian distribution of power (Fernandez and Rodrik 1991). If winners and losers cannot be identified *ex ante*, all agents in society might oppose a reform that might have beneficial net effects. If the distribution of power is less equal, the problem becomes more severe as the agents in privileged positions might fear a double loss from reform: the direct loss (for instance, in terms of rents extracted), and the indirect loss of power that might follow from reform.

The details of the design of institutions, understood as the set of rules that mediate between the preferences of social actors and the outcomes, affect the outcomes themselves. In general, one can claim that the way in which institutions are designed reflects in itself the preoccupation, by the 'winners' of today, to secure victories in the future. This general approach can guide the way in which one can think of governance. We propose to understand 'bad governance' in this light. By 'bad' (or sub-optimal) governance we mean those policies that fail to put a given society as close as possible to its 'prosperity potential'. By 'possible' we mean *after taking into account* the fact that some policies, albeit leading to desirable outcomes, are not implementable because of conflicts with basic principles, or because of unintended consequences that are built in the way in which they can be achieved.

Understanding the causes of bad policy from a political economy point of view is more puzzling than it might seem in the layperson's eyes. Indeed, as laid out in a seminal paper, 'Theories of "Bad Policy"', by James Robinson (1998), then developed into several papers in Acemoğlu et al. (2001), standard explanations of inefficient policies often do not survive simple analytical scrutiny. On the one hand, inefficient policy might be simply the consequence of poor implementation (as in the perspective of Evans 1992, 1995; Evans and Rauch 1999; or Huber and McCarty, 2004). On the other hand, 'bad policy', in the sense of policy designed so that its consequences, regardless of the quality of the implementation, are inefficient, requires explanations that are more complex than what one might think. Indeed, neither theories that posit 'government as a veil'—i.e. where government action simply reflects the balance of power of different groups in society—nor those that assume state autonomy, and in particular the 'predatory state' approach (e.g. Buchanan and Tullock 1962), can explain why inefficient policies are designed.

In the 'predatory state' approach, for example, the government apparatus is conceived as populated by self-interested agents that use the machinery of the state to extract resources from society and appropriate/steal them. However,

Piero Stanig and Mark Kayser

one would suppose that even the most predatory dictator would have an interest in implementing efficient policies, simply for the reason that there would be more for him to prey on if efficient policies are pursued. This is the intuition behind Mancur Olson's (2000) 'stationary bandit' analogy: as long as the 'bandit' has a long-term interest in extracting resources from a given society, it has an interest in providing those public goods that make that same society produce as much as it can. In other words, why not maximise the size of the pie one wants to steal, and focus on stealing later?

Yet, what can be called 'bad policy' is a provision that reduces the size of the pie that is going to be stolen. And a set of phenomena labelled as 'governance failures' cannot be ascribed simply to poor implementation: they often seem to be inefficient by design. Bad policies that reduce the size of the pie on purpose happen both in ideal-typical dictatorships, where a small group of people (e.g. the dictator's inner circle, or single party leaders) might keep a society well below its production frontier, or, in other words, below what would be a society's 'prosperity potential', as well as in textbook democracies, in which organised 'special interests' exert pressure on decision-making bodies to seek rents at the expense of efficiency (Grossman and Helpman 2001) or in which a median voter prefers more redistribution than the level that would maximise income growth (Alesina and Rodrik 1994).

As in the classic logic of 'rent seeking', the distributive concern of specific groups or individuals in society trumps efficiency: some prefer a larger share of a smaller pie, and resources are used unproductively to wrestle existing wealth, rather than producing more of it. Robinson's (1998) intuition is that inefficient (or unequivocally 'bad') policies are implemented exactly because they preserve a given distribution of political power. 'Good' (or efficient) policies might have, among their consequences, that of altering the balance of power in a given society.

It is worth keeping in mind that good governance is not necessarily Pareto-improving. In other words, it is not always the case that everyone benefits from a move from bad to good governance, even if the net benefit to society of good governance is positive. For a simple example, think about a policy that reduces petty corruption: while improving quality of service provision and reducing the cost paid by the 'general public' (or users of a given government service), at a minimum it reduces the standard of living of the corrupt officials! Similarly, a policy that allocates benefits based on merit negatively affects those who lack the merit but would be granted those benefits if resource allocation were based, for instance, on 'connections'. In general, a move from discrimination or partiality to impartiality cannot be a Pareto improvement, as all those who benefitted from the discriminatory allocation are going to be worse off, unless the efficiency gains from the new, impartial allocation are sufficiently large as to compensate the former winners from

204

6/6/2013   12:53:38 PM

discrimination (that become, then, losers from the reform and from impartial allocation).

An important dimension of governance, then, has to do with the extent to which power, even in a democratic setting, is in the hands of social actors in privileged positions who might want to prevent any change in the allocation of power, as opposed to the 'median voter' or a set of fluid coalitions between social actors drawn from different locations in the income distribution and different positions in the productive system. Often a major concern for those who occupy political office is to deliver benefits to powerful economic interests in exchange for promises of present or future rewards. Here, we set out to measure one specific facet of this phenomenon, looking at the extent to which ties between holders of political offices and large corporations might bias high-level policymaking in the direction of (a) favouring those large corporations, or (b) ensuring that those same powerful corporations remain profitable over time.

Political capture: Links between government officials and private interests

Legal activities that distort the decision making of public bodies for private gain pose an important problem for effective governance. Few comparative indicators exist that measure the extent to which private interests guide public policy. This indicator addresses a fundamental question: For whom does the government govern?

One can argue that most misgovernance is a consequence of privileged interests persuading the state to govern in their interest, possibly getting their way against the interest of the broader public. Such forms of 'capture' are not necessarily illegal. Many countries impose some limits on conflict of interest that might result from appointments of former office holders to corporate office. Nevertheless, even when they respect such rules, connections between special interests and office holders might damage the well-being of the governed and weaken the stewardship of the state in question.

When government action is driven by the priority to cater to the needs of well-placed actors and representatives of special interests, rather than to those of the majority of citizens, its ability to implement policies that respond to policy challenges is diminished. Special interests from their position of privilege might be able to keep off the agenda policy changes that could erode their privilege itself; or they might alter the course of debates and restrict the options discussed. A captured government is inherently slow-reacting and conservative, in the specific meaning of resistant to change and opposed to innovation, because it is not in the interest of the

Piero Stanig and Mark Kayser

well-positioned (rent seekers) to alter arrangements that sustain their ability to extract rents.

Suboptimal policies manifest themselves in different sectors with different actors in different states, organisations, and contexts, be it preferences for the security of financial institutions over job growth in Europe, subsidies to non-competitive agriculture in Japan, or giveaways from the public purse to energy firms and the emasculation of environmental and safety regulation in the US. When the privileged can influence the rules, they do so in their favour. How this varies in degree and manner is a matter for measurement.

Our interest lies in general on the causes and effects of capture on policy-making, and more specifically in 'revolving door' mechanisms. A cross-national measure capturing the revolving door is particularly useful, we believe. Several studies have been carried out to explore the potential effects of 'capture' in general, and of 'revolving doors' phenomena more specifically. Most of them, though, have a relatively narrow focus, looking at regulatory practice within single countries, and in particular in the United States. Our interest, by contrast, lies in the general effects of capture on policymaking, in a cross-national perspective.

Studies show that regulation seems to be affected, across states in the US, by campaign contributions by firms (De Figueiredo and Edwards 2007); moreover, the behaviour of regulators is associated with their own background in a given industry (Gormley 1979; Cohen 1986) and to the mechanism by which they are chosen, appointed, or elected (Boyes and Dowell 1989; Kwoka 2002; Besley and Coate 2003). Detailed descriptive accounts of the revolving door phenomenon have been done for the United States. Etzion and Davis (2008) document how members of the higher rungs of the executive under Clinton and G. W. Bush moved to corporate board appointments: for instance, all the five members of the Joint Chiefs of Staff in 2002 were appointed to corporate boards of 'central firms in the defense industry' (161).

In introducing their study on the financial returns of office-holding in the British Parliament, Eggers and Hainmueller (2009: 513) note the strong relationship between business and government the world over:

Firms with personal and/or financial connections to politicians have enjoyed higher stock valuations in Indonesia (Fisman 2001), the United States (Goldman, Rocholl, and So, n.d. [2009]; Jayachandran 2006; Roberts 1990), Malaysia (Johnson and Mitton 2003), and Nazi Germany (Ferguson and Voth 2008). In the United States, politically connected firms are more likely to secure procurement contracts (Goldman, Rocholl, and So 2008 [2010]), and in Pakistan, they are able to draw more favorable loans from government banks (Khwaja and

Mian 2005). Faccio (2006) shows that the benefits of political connections are larger in countries with higher corruption scores.

In a study that aims at detecting causal effects, Blanes i Vidal et al. (2010) estimate the substantial premium that former government officials who work as lobbyists in the US receive from their connections to politicians. For instance, lobbyists who worked with a given senator see their revenues drop on average by 21 per cent when that senator leaves office. Needless to say, the fact that lobbyists with direct ties to a high-level decision maker are considered more valuable by their customers implies that they can provide them with a more valuable service in terms of influence on policy than those who lack such direct connections.

There is also evidence that serving in the British Parliament triples the probability that a Conservative politician later serves as a director of a publicly traded firm and nearly doubles his/her wealth (Eggers and Hainmueller 2009). As the authors suggest, the difference between the benefits enjoyed by the Conservative MPs and the Labour MPs can be explained by the fact that the latter, because of party discipline and oversight on the part of trade unions, are not as able as the Conservative MPs to sell their services to private corporations.

Cross-national analysis is lacking, and one can attribute the underdevelopment of empirical research on this topic to the absence of comparable cross-national assessments of the extent of revolving door phenomena and of the degree of capture of policymaking by the 'well-positioned'. One cross-national study uses scores of *corruption* as a proxy for how 'capturable' regulators are to show that, across Latin America, when and where regulators are easier to capture, electricity suppliers are less efficient (Dal Bó and Rossi 2004).

In a recent study, Kaufmann and Vicente (2011) use data from a cross-national survey of firms (the Executive Opinion Survey by the World Economic Forum) to analyse what institutional and structural characteristics of countries are associated with the prevalence of what they call 'legal' corruption, and we call 'capture', relative to traditional 'illegal' corruption. Kaufmann and Vicente (2011) measure perceptions of 'legal' corruption with survey items about the role of influence of the well-connected in procurement, the role of legal contributions to political parties, the independence of the judiciary from influence, and influence on laws and regulations. The authors limit themselves to exploring the possible causes of the prevalence of legal corruption, underlining especially the role of accountability in the control of the forms of 'legal' corruption that are prevalent in more developed and more equal countries. The authors also suggest that understanding the consequences of legal corruption on economic performance should be the next item in the research agenda.

Piero Stanig and Mark Kayser

We propose here a cross-national measure of governmental capture based on objective measures of the frequency in which high governmental officials can be found on the boards of major corporations following the end of their employment in government.[6] Our revolving doors measure adopts an approach similar to that of Faccio (2006) to estimate the political connections of firms. Where Faccio measures connections from corporate boards to parliament and government in order to examine benefits to politically connected firms, we focus our attention first on government officials to determine how many of them later enjoy employment with private sector firms. Many corporations, lobbying firms, public relations firms, and law firms recruit government officials for post-governmental employment both to influence their decisions while still in government and to influence the decisions of former colleagues after leaving government. Collecting such data internationally is difficult. Nevertheless, the revolving doors indicator will measure the proportion of high officials in finance, economic, and defence ministries who find employment in large corporations after leaving office.

Proposal 2: Innovation Capacity in a Societal Context

Organisational 'capture', as important as it is, has no monopoly on under-represented measures of governance. To this list we must add a diverse set of measures that capture aspects of governance such as capacity and infrastructure for policy innovation that emerge from the interconnectedness of governing structures and actors, political and institutional structures, and third sector non-governmental resources. Governance rests on a broad foundation.

In contrast to our proposal for a state capture indicator in the section above, here we judge this aspect of governance to be best depicted by a dashboard of indicators. In particular, we look at those resources located outside of government which might complement those in the government bureaucracy itself. We approach this question from multiple perspectives. On the one hand, we look at resources that might have a crucial role in the generation of governance innovation: think tanks, academic researchers, and policy schools. On the other, we collect information about the advocacy infrastructure in the country.

---

[6] We have developed and continue to work on this idea, together with Klaus Broesamle. For a more indepth look at 'capture' and a more detailed proposal for collecting data on the phenomenon, see Broesamle's chapter in this volume.

We start from the premise that a dense and egalitarian civil society contributes to good governance from different sides. First of all, it articulates social demands and gathers information about problems that might not be optimally solved at a given moment. In addition, it monitors policy implementation and prevents opportunistic behaviour on the part of the government. Finally, it might directly contribute to policy implementation in the context of public/private partnerships. The egalitarian dimension of civic involvement is something we consider particularly important. If only the privileged (for instance, the wealthiest, or the most educated) in society are engaged in civil society, the demands put forward and the interests represented might provide a biased picture of the actual demands and interests.

Before describing the more innovative/novel aspects of our contribution, we need to clarify the link between knowledge and governance. Poor governance might derive not from purposefully inefficient design, but simply from a lack of the administrative and technical resources needed to design and implement policies. In this case, poor governance does not have to do with inefficient policy goals, as in the case of capture, but with ineffective policy design and implementation when well-intentioned policymakers lack expertise and resources. Effective policies, including those that require coercive enforcement and therefore the intervention of the state, often require not only competent policymakers but knowledge resources, both inside and outside of government. Indeed, effective and innovative policy proposals cannot be expected to come exclusively from within the state apparatus itself.

We propose a set of measures intended to assess the extent of intellectual resources available to governments needed to diagnose problems, explore avenues of innovation, and design context-appropriate policies. We draw on numerous measures of bureaucratic capacity, external policy expertise, and social capital/civic engagement to propose a dashboard of innovation capacity.

We address each dimension separately because they capture separate and substantively meaningful (and potentially orthogonal) constructs: in other words, some governments might display high capacity, but there might be few external knowledge resources for them to draw on, and civil society might be weak; and, conversely, a society with significant knowledge resources and an active and egalitarian civil society might face a low-capacity bureaucracy. Learning separately about each of the dimensions is therefore useful.

Bureaucratic capacity

Since Max Weber's classic work on bureaucracy, conventional wisdom has held the idea that the ability of government agencies to implement policy

Piero Stanig and Mark Kayser

depends on the recruitment practices and internal cohesion of the bureau-
cratic apparatus. In more recent times, scholars (Evans and Rauch 1999) have
proposed a more theory-rich way of thinking about the bureaucracy, specify-
ing the causal mechanisms underlying the link between design of govern-
ment agencies and policy outcomes. In other words, they attempt to open
the 'black box' of governance that, more or less deliberately, characterises
research that focuses on bureaucratic corruption and bureaucratic red tape
and delay.

Evans and Rauch (2000) summarise two perspectives—one more sociologi-
cal, one more economic in approach—that explain the relationship between
recruitment and promotion practices and good governance outcomes. In
the sociological view laid out by Evans (1992) and Evans (1995), Weberian
selection and promotion of bureaucrats lead to the selection of 'a capable
pool of officials'; moreover, rules for career advancement within agencies
foster the formation of ties between bureaucrats and a concern about reputa-
tion among officials that promotes compliance with formal rules as a conse-
quence of the development of a 'commitment to corporate goals' backed by
an *esprit de corps*.

In the view of Rauch (1995), internal promotion plays a bigger role than
salaries and initial recruitment. By promoting to supervisory roles those
employees that are more interested in power, understood as the ability to
implement their preferred policy objectives, 'internal promotion generates a
virtuous circle that increases (in expectation) the value the principal places
on exercise of power, tending to increase the extent to which the bureauc-
racy as a whole carries out its assigned tasks…and decrease the extent to
which it implicitly taxes the private sector through large-scale corruption'
(Rauch and Evans 2000: 53). In other words, ambition leads officials to act as
strict supervisors of their supervisees. On the one hand, more power-hungry
officials monitor their employees more closely; on the other, they are more
likely to ascend in the agency's hierarchy.

In the model proposed by Huber and McCarty (2004), the direct effects
of low bureaucratic capacity are twofold: first, policy does not address an
issue effectively because of mis-implementation of policy directives; sec-
ond, bureaucrats with less capacity are less likely to comply with the pol-
icy directives. Less able bureaucrats, in their model, are more likely to
implement policies incorrectly. Knowing that, by mistake, they might vio-
late the directives coming from politicians (because the policy the bureau-
crats implement is not what the bureaucrats themselves intended to
implement), bureaucrats also have fewer reasons to try to follow directives
and implement policies within the boundaries set by politicians. After all,
the value to them of an extra effort to comply with directives is small,
because their attempts to comply with directives might be frustrated by

their 'technical' inability to actually implement the policy they had intended to implement.[7]

External know-how and innovation capacity

So how have bureaucratic quality and know-how been measured? And how should they be measured? First, consider bureaucratic quality. We believe that limiting oneself to measurement of the quality of the bureaucracy—for instance, via measurement of the degree to which recruitment is merito-cratic (Evans and Rauch 1999; Dahlström et al. 2011), or via indices that capture prevalence of corruption (Transparency International) or bureau-cratic delays (World Bank Doing Business Indicators), leaves out an impor-tant part of *governance*: namely, the ability to respond effectively to policy challenges.

Not only do the direct implementers need to have capacity, but technical and intellectual resources are needed. For instance, policy proposals, even in their details, might require expertise that exceeds the resources available in a given agency. Moreover, closed recruitment and internal promotion, a com-ponent of the Weberian model, might lead as a by-product to a homogenous culture within the bureaucracy and an over-reliance on tried-and-tested approaches to deal with the external world. A negative consequence of the Weberian selection process, then, might be the inability of an agency to tackle new problems and face new challenges that might emerge. A conse-quence of a lack of intellectual 'cross-pollination', and criticism of estab-lished practices to deal with problems, is intellectual sclerosis. The availability in a given society of alternative proposals regarding ways to deal with given issues and the existence of internationally recognised specialists in academia affect, to a large extent, policy design and implementation in the face of unexpected or new challenges.

Our definitions encompass both what Huber and McCarty (2004) call 'expertise'—i.e. the ability to design policy so that it has the desired effect,

---

[7] Huber and McCarty (2004) also look at the spillover effects from bureaucratic capacity, or its absence. First of all, given that incompetent bureaucrats are harder to make comply with direc-tives, politicians have more incentives to hire bureaucrats that are politically close to them, lead-ing to the politicisation of the bureaucracy. Moreover, when capacity is low, politicians have little incentive to reform other parts of the government, like the judiciary, that can be used to monitor the behaviour of bureaucrats. Indeed, many instances of mis-implementation are due to mere mistake, and the threat of prosecution does not affect this type of mistake, if the bureaucrat is simply incapable. On the other hand, if politicians are capable, the threat of prosecution for mis-implementation affects more the behaviour of the bureaucrats, inducing them to comply with policy directives coming from above. Hence, a functioning judiciary is much more valuable as a monitoring device for capable bureaucrats. Countries might get stuck in a no-reform, low-capac-ity equilibrium in which increases in monitoring do not seem to be worth their cost given the small effects they would have on compliance, and increases in capacity do not seem to be worth their cost given the weak monitoring devices at the politician's disposal.

Piero Stanig and Mark Kayser

given the information one can gather about the state of the world; and what they call 'capacity'—i.e. the ability, once a policy is designed, to implement it in the way it is intended. The governmental and societal resources that we try to capture in our dashboard contribute to both of these dimensions. Our approach goes beyond bureaucratic capacity and involves the availability of intellectual resources and know-how in a given society and in and outside of government.

Policy is not necessarily developed in isolation from within bureaucratic agencies, and it is not implemented in a vacuum. High-level bureaucrats, as well as ministers and junior ministers that decide the overall direction of policies, are often recruited internally. We believe that, to assess the ability of governments to implement policy that responds to challenges, both traditional/old and new and unexpected, and possibly in innovative ways, depends also on the amount of resources available in a given society in terms of vibrancy of the 'marketplace of ideas'. Our dashboard on innovation capacity recognises this point and tries to provide ways to assess not only the professionalism of the bureaucratic apparatus in the standard Weberian tradition, but also the diffuse resources that are available for policy design and governance innovation. Both think tanks, which generate and advance innovative policies by providing fact-based advice regarding their costs and benefits, and advocacy groups associated with a robust civil society, contribute to policy innovation.

The innovation capacity dashboard

While this chapter focuses on the theoretical and methodological foundations needed for better governance indicators, a companion publication (Anheier et al. 2013) dedicates itself to actually implementing some of them. A few notable differences appear—as can be expected between projects focused, respectively, on theory and implementation. The applied project also differs substantively by collecting data on civil society strength, that we do not discuss here, and by not including, as of yet, the revolving door indicator of state capture. Indeed, the components of the present innovation capacity dashboard are a subset of the National Governance Dashboard in Anheier et al. (2013). We prioritise the narrower dashboard here for its more obvious single dimensionality and concordance with the theory of indicator design emphasised in this chapter. Despite such divergences, the theoretical and methodological case laid out here very much underpins the indicator development of that project and, we hope, other future projects farther afield and unrelated to the present authors as well.

The remainder of this section addresses the resources that affect the success of innovation capacity: the ability, once a policy is designed, to implement

212

it in the way it is intended (implementation capacity) and the ability to design policy so that it has the desired effect, given the information one can gather about the state of the world (expertise). Although we do not develop any indicators here, we remain as concrete as possible—and even identify data sources—so that this chapter can serve as a road map for developing the indicators we propose.

The innovation capacity dashboard captures two key dimensions of governance: internal governmental implementation capacity and external expertise, as described below.

IMPLEMENTATION CAPACITY

Implementation capacity is measured as (a) the Weberianness of a bureaucracy, i.e. its impartiality and professionalism in hiring and promotion; (b) the statistical capacity, i.e. the ability to diagnose problems through the collection of social and economic data; and (c) the intellectual resources within government, measured by the number of researchers with an advanced degree employed by the government.

Data from the Quality of Government (QoG) Institute (Teorell et al. 2011) that address the issue of Weberianness, with a survey of experts in a large set of countries, offer a cornerstone for this dashboard. The survey was carried out between 2008 and 2010. The authors of the report summarise the answers to several questions in three indices, of which two pertain to our purposes: the index of bureaucratic impartiality, and the index of bureaucratic professionalism. The index of impartiality measures to what extent government institutions exercise their power impartially. The impartiality norm is defined as follows: 'When implementing laws and policies, government officials shall not take into consideration anything about the citizen/case that is not beforehand stipulated in the policy or the law' (Rothstein and Teorell 2008: 170). The index of professionalism measures to what extent the public administration is professional, rather than politicised (Dahlström et al. 2011). Higher values indicate a more professionalised public administration.

As we briefly sketched in our review of the Huber and McCarty (2004) framework above, policymaking requires knowledge of the 'state of the world' to which policy is responding. Without the ability to measure actual conditions on the ground, policy cannot be designed to produce the intended outcomes. Thus, the index of statistical capacity compiled by the World Bank for over 140 developing countries is of great value for measuring this aspect in at least those countries. Developed countries, unsurprisingly, generally have high and less varied statistical capacity. Using information available from the International Monetary Fund, United Nations, UNESCO, and the World Health Organisation and its own information, the World Bank scores a country against specific criteria along three dimensions (statistical

Piero Stanig and Mark Kayser

methodology, source data, and periodicity and timeliness) and derives an overall score for each country on a scale of 0–100, with a score of 100 indicating that the country meets all the criteria.

To round out the measurement, one can capture the amount of intellectual capacity available within the public administration, albeit roughly, with two variables collected for several countries by UNESCO: the (log) full-time equivalent (FTE) number of holders of advanced degrees (in all fields) employed by the government, and the (log) FTE number of social science researchers employed by the government.

EXPERTISE

Governance extends beyond the capacities of government. As we discussed above, external know-how and capacity for innovation can be central to designing innovative and context-specific policies. One can capture the extent of such external sources of knowledge and expertise with data that measure think tanks, top economics departments, policy schools, the number of social science researchers, and gross domestic expenditure for the social sciences as a proportion of total research and development expenditures.

To address the role of think tanks, we identify three different sources. The first is the Global Go To Think Tanks Report. The Think Tanks and Civil Societies Program at the University of Pennsylvania collects data on think tanks and reports, yearly; a set of rankings, at the global level, by region, and by issue area. It also reports the total number of think tanks present in each country according to the information available to the organisation. Information, both from the rankings (that reflect quality of the think tanks in a given country) and the counts (that reflect quantity irrespective of quality), are relevant. The second source is IDEAS RePEc, a repository for academic papers. It reports a ranking of the top 25 per cent of the think tanks with members registered to the repository based on their scientific output. The third is the National Institute for Research Advancement (NIRA)[8], an organisation based in Japan that keeps an updated list of think tanks around the world.

Data on academic resources are scattered over different sources. The ranking of economics departments at a global and regional level is the first source (IDEAS 2011). While clearly these rankings can be found lacking from the point of view of completeness, given that they are based on self-selected authors in the most popular article repository, at a minimum they measure the presence of academic economics departments whose members participate in the international scientific community. Scholars who are successful internationally are more likely to command the methodological sophistication necessary to publish in prominent journals and convince others that

[8] <http://www.nira.or.jp/>

their models and estimates are accurate. Academically successful economists are also more likely to stay on top of the developments in the field and produce innovative ideas. In fact, any of several disciplines in the social sciences could be used for this indicator, but the compulsion of economists for self-ranking greatly simplifies data collection.

In order to assess the quality and vitality of the academic fields that mostly affect proposals for innovation and responses to new challenges in governance, we also identify (along with the data about ranking of academic departments of economics) data published by UNESCO on the full-time equivalent (FTE) number of researchers in a given country (both in general, and specifically in the social sciences) per million inhabitants, and gross domestic expenditure for research (again, in general, and for the social sciences) per million US dollars GDP. Finally, the (log) count of number of policy graduate programmes in each country, as listed in the website GradSchools.com,[9] serves as a rough measure of bureaucratic training. While this constitutes a third-party count (unverified by us) without a ranking, and therefore can be treated as preliminary data only, it is worth stating that (a) there is no ranking of policy schools outside of the United States, and (b) this is a first step towards coming up with better measures of the quality of professional higher education for policymakers around the world.[10]

The variables and data sources identified here provide the components for the construction of a dashboard of policy innovation (and implementation) capacity that together with measure of state capture (like that outlined in the previous section) address important gaps in current efforts to measure governance. Both the revolving doors indicator of state capture and the innovation capacity dashboard connect government to some of the key societal actors that influence the success or failure of policy. Measuring governance in the context of actors beyond government, we argue, is the key progress in the development of indicators that capture the effectiveness and efficiency of governance.

Conclusion

The measurement of governance has become an increasingly crowded field over the last few decades. It is nevertheless, as we argue above, a field that

---

[9] <www.gradschools.com/>.

[10] When dealing with rankings of think tanks and economics departments, we adopt an approach similar to Borda scoring. Namely, if K organisations are ranked, we assign the score K+1 to the top ranked organisation, K to the second organisation, etc. We then sum the ranks for each country, and we assign the score of 1 to countries that do not have any organisation appearing in the ranking.

Piero Stanig and Mark Kayser

neglects important aspects of governance that are essential to the generation of effective and efficient policy. While many of the better cross-national measures focus on the control of corruption or bureaucratic efficiency, very few governance indicators capture the central determinants of bad policy—whether the ability of privileged interests to skew and maintain public policy in their favour, or the incapacity of government to analyse and innovate.

Governance demands the capacity to respond to challenges through policy innovation. Organisations captured by narrow interests, however, cannot respond as well or as quickly when captured policymakers preclude policy options that contravene the specific interests of a few. Policymaking can also become ossified when those who benefit from established policies see all change as a risk to a preferred status quo. The privatisation of public policy is a key determinant of poor policy, but also one that, until now, has gone largely unmeasured.

Governance is also less effective where governments do not have the resources to diagnose problems correctly, to generate innovative context-specific solutions, and to implement them faithfully. Governance obviously depends on intellectual and bureaucratic capacity. But the recognition of new challenges and the development and implementation of solutions emerges, increasingly frequently, from organisations outside of government. Governance rests and relies on a societal foundation. Where societies and civil society are robust, citizens are more able to organise to advocate for their interests and more likely to abide by governance decisions. Good governance depends on the governed, and the governed are organised to identify and advocate for their interests through civil society.

Both of the sets of indicators we propose here place government in a societal context. Rather than atomistic measures of governmental structure—'government indicators'—we call for and propose 'governance indicators' which measure the performance of governance in connection with business and societal interests. Designing effective, efficient, and responsive governance will always be a difficult and politicised task as long as policymaking has distributional consequences. Understanding how policies can deliver the most benefit to society at the least cost, and identifying which systems work well and which do not, is best done through comparison. By developing new cross-national measures for key aspects of governance we hope to enable such comparison and efforts at improvement. This study and these indicators are only the first step in what we intend to be a long effort in the study of governance. We look forward to the continued process of designing and collecting governance measures that build on the foundation we have set here.

References

Acemoğlu, D., Johnson, S., and Robinson, J. A. (2001). 'The Colonial Origins of Comparative Development: An Empirical Investigation', *American Economic Review*, 91: 1369–401.

Alesina, A. and Rodrik, D. (1994). 'Distributive Politics and Economic Growth', *The Quarterly Journal of Economics*, 109(2): 465–90.

Anheier, H. K., Stanig, P., and Kayser, M. (2013). 'Introducing a New Generation of Governance Indicators', in Hertie School of Governance (ed.), *The Governance Report 2013*. Oxford: Oxford University Press, 117–48.

Bajada, C. and Schneider, F. (2005). 'The Shadow Economies of the Asia-Pacific', *Pacific Economic Review*, 10(3): 379–401.

Bandura, R. (forthcoming). *Composite Indicators and Rankings: Inventory 2011*. Brussels: European Commission Joint Research Centre.

Besley, T. and Coate, S. (2003). 'Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach', *Journal of Public Economics*, 87(12): 2611–37.

Blanes i Vidal, J., Draca, M., and Fons-Rosen, C. (2010). *Revolving Door Lobbyists*, CEP Discussion Paper No. 993. London: Centre for Economic Performance, London School of Economics and Political Science.

Bollen, K. A. (2002). 'Latent Variables in Psychology and the Social Sciences', *Annual Review of Psychology*, 53: 605–34.

Boyes, W. J. and McDowell, J. M. (1989). 'The Selection of Public Utility Commissioners: A Reexamination of the Importance of Institutional Setting', *Public Choice*, 61(1): 1–13.

Breusch, T. (2005). *Estimating the Shadow Economy using MIMIC Models*, Working Paper. Canberra, Australia: The Australian National University.

Buchanan, J. M. and Tullock, G. (1962). *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor, MI: University of Michigan Press.

Cohen, J. E. (1986). 'The Dynamics of the "Revolving Door" on the FCC', *American Journal of Political Science*, 30(4): 680–708.

Dahlström, C., Lapuente, V., and Teorell, J. (2011). 'Dimensions of Bureaucracy II: A Cross-national Dataset on the Structure and Behaviour of Public Administration', QoG Working Paper Series 6. Gothenburg: The Quality of Government Institute, University of Gothenburg.

Dal Bó, E. and Rossi, M. A. (2004). 'Corruption and Inefficiency: Theory and Evidence from Electric Utilities', Working Paper, <http://www.udesa.edu.ar/files/UAEconomia/Seminarios%20y%20Actividades/Seminarios%20Permanentes/2004/rossi.pdf> (accessed on 26 March 2012).

De Figueiredo, R. J. P. and Edwards, G. (2007). 'Does Private Money Buy Public Policy? Campaign Contributions and Regulatory Outcomes in Telecommunications', *Journal of Economics and Management Strategy*, 16(3): 547–76.

Dell'Anno, R. and Schneider, F. (2006). *Estimating the Underground Economy by Using MIMIC Models: A Response to T. Breusch's Critique*, Economics Working Papers 2006–7. Linz, Austria: Department of Economics, Johannes Kepler University of Linz.

Piero Stanig and Mark Kayser

Digman, J. M. (1990). 'Personality structure: Emergence of the five-factor model', *Annual Review of Psychology*, 41: 417–40.

Djankov, S., La Porta, R., López-de-Silanes, F., and Shleifer, A. (2002). *Courts: The Lex Mundi Project*, CEPR Discussion Papers 3344. London: CEPR.

Eggers, A. C. and Hainmueller, J. (2009). 'MPs for Sale? Returns to Office in Postwar British Politics', *American Political Science Review*, 103(4): 513–33.

Etzion, D. and Davis, G. F. (2008). 'Revolving Doors? A Network Analysis of Corporate Officers and U.S. Government Officials', *Journal of Management Inquiry*, 17(3): 157–61.

Evans, P. B. (1992). 'The State as Problem and Solution: Predation, Embedded Autonomy, and Adjustment', in S. Haggard and R. R. Kaufman (eds.), *The Politics of Economic Adjustment*, Princeton, NJ: Princeton University Press, 139–81.

Evans, P. B. (1995). *Embedded Autonomy: States and Industrial Transformation.* Princeton, NJ: Princeton University Press.

Evans, P. and Rauch, J. E. (1999). 'Bureaucracy and Growth: A Cross-National Analysis of the Effects of "Weberian" State Structures on Economic Growth', *American Sociological Review*, 64(5): 748–65.

Evans, P. and Rauch, J. E. (2000). 'Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries', *Journal of Public Economics*, 75(1): 49–71.

Faccio, M. (2006). 'Politically Connected Firms', *American Economic Review*, 96(1): 369–86.

Ferguson, T. and Voth, H.-J. (2008). 'Betting on Hitler—The Value of Political Connections in Nazi Germany', *The Quarterly Journal of Economics*, 123(1): 101–37.

Fernandez, R. and Rodrik, D. (1991). 'Resistance to Reform: Status Quo Bias in the Presence of Individual–Specific Uncertainty', *American Economic Review*, 81(5): 1146–55.

Fisman, R. (2001). 'Estimating the Value of Political Connections', *American Economic Review*, 91(4): 1095–102.

Giles, D. E. A. and Tedds, L. M. (2002). *Taxes and the Canadian Shadow Economy*. Toronto: Canadian Tax Foundation.

'The Global Go To Think Tanks Report' (2011). *Think Tanks and Civil Society.* Retrieved from <http://www.gotothinktank.com/thinktank/> (accessed on 16 March 2012).

Goldman, E., Rocholl, J., and So, J. (2009). 'Do Politically Connected Boards Affect Firm Value', *Review of Financial Studies*, 22(6) 2331–60.

Goldman, E., Rocholl, J., and So, J. (2010). *Political Connections and the Allocation of Procurement Contracts*, Working Paper, <http://www.istfin.eco.usi.ch/j_rocholl.pdf> (accessed 16 March 2012).

Gormley, W. T. (1979). 'A Test of the Revolving Door Hypothesis at the FCC', *American Journal of Political Science*, 23(4): 665–83.

Grossman, G. M. and Helpman, E. (2001). *Special Interest Politics*. Cambridge, MA and London: MIT Press.

Hagerty, M. R. and Land, K. C. (2007). 'Constructing Summary Indices of Quality of Life: A Model for the Effect of Heterogeneous Importance Weights', *Sociological Methods and Research*, 35(4): 455–96.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edn. New York: Springer Verlag.

Huber, J. D. and McCarty, N. (2004). 'Bureaucratic Capacity, Delegation, and Political Reform', *American Political Science Review*, 98(3): 481–94.

IDEAS (2011). 'Research Papers in Economics Database (RePEc)', <http://ideas.repec.org/> (accessed on 16 March 2012).

Jayachandran, S. (2006). 'The Jeffords Effect', *Journal of Law and Economics*, 49(2): 397–425.

Johnson, S. and Mitton, T. (2003). 'Cronyism and Capital Controls: Evidence from Malaysia', *Journal of Financial Economics*, 67(3): 351–82.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). 'Response to What do the Worldwide Governance Indicators Measure?', *European Journal of Development Research*, 22(1): 55–8.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). 'The Worldwide Governance Indicators: Methodology and Analytical Issues', *The Hague Journal on the Rule of Law*, 3(2): 349–79.

Kaufmann, D. and Vicente, P. C. (2011). 'Legal Corruption', *Economics & Politics*, 23(2): 195–219.

Khwaja, A. J. and Mian, A. (2005). 'Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market', *The Quarterly Journal of Economics*, 120(4): 1371–411.

Knack, S. and Keefer, P. (1995). 'Institutions and Economic Performance: Cross-country Tests Using Alternative Institutional Measures', *Economics & Politics*, 7(3): 207–27.

Kwoka, J. E. (2002). 'Governance Alternatives and Pricing in the U.S. Electric Power Industry', *Journal of Law and Economic Organization*, 18(1): 278–94.

Mauro, P. (1995). 'Corruption and Growth', *Quarterly Journal of Economics,* 110(3): 681–712.

Olson, M. (2000). *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. New York: Basic Books.

Rauch, J. E. (1995). 'Bureaucracy, Infrastructure, and Economic Growth: Evidence from U.S. Cities During the Progressive Era', *The American Economic Review*, 85(4): 968–79.

Rauch, J. E. and Evans, P. (2000). 'Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries', *Journal of Public Economics*, 75(1): 49–71.

Roberts, B. E. (1990). 'A Dead Senator Tells No Lies: Seniority and the Distribution of Federal Benefits', *American Journal of Political Science*, 34(1): 31–58.

Robinson, J. A. (1998). 'Theories of Bad Policy', *Policy Reform*, 1(1): 1–46.

Rothstein, B. and Teorell, J. (2008). 'What is Quality of Government? A Theory of Impartial Government Institutions', *Governance*, 21(2): 165–90.

Schwartz, S. H. (1992). 'Universals in the Content and Structure of Values: Theory and Empirical Tests in 20 Countries', in M. Zanna (ed.), *Advances in Experimental Social Psychology*. New York: Academic Press, 1–65.

Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris: Commission on the Measurement of Economic Performance and Social Progress.

Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2010). *Mismeasuring Our Lives: Why GDP Doesn't Add Up*. New York: The New Press.

Piero Stanig and Mark Kayser

Teorell, J., Samanni, M., Holmberg, S., and Rothstein, B. (2011). 'The Quality of Government Basic Dataset, made from the QoG Standard Dataset version 6 April 2011'. Gothenburg: The Quality of Government Institute, University of Gothenburg, <http://www.qog.pol.gu.se> (accessed on 16 March 2012).

Thomas, M. A. (2010). 'What Do the Worldwide Governance Indicators Measure?', *European Journal of Development Research*, 22(1): 31–54.

Transparency International (TI) (2011). 'Corruption Perception Index', <http://cpi.transparency.org/cpi2011/> (accessed on 16 March 2012).

UIS (UNESCO Institute of Statistics) (2012). 'Data Centre', <http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=136&IF_Language=eng&BR_Topic=0> (accessed on 16 March 2012).

UNDP (United Nations Development Programme) (2004). *Governance Indicators: A Users' Guide*, 2nd edn. Oslo: Oslo Governance Centre.

World Bank (2012). 'Doing Business', <http://www.doingbusiness.org> (accessed on 16 March 2012).