

# Delivering Service Indicators in Education and Health in Africa

## A Proposal

*Tessa Bold*

*Bernard Gauthier*

*Jakob Svensson*

*Waly Wane*

The World Bank  
Africa Region  
Poverty Reduction and Economic Management



Middle East and North Africa Region  
Public Sector and Governance  
June 2010

## Abstract

The Delivering Service Indicators seek to provide a set of indices for benchmarking service delivery performance in education and health in Africa in order to track progress in and across countries over time. It seeks to enhance effective and active monitoring of service delivery systems and to become an instrument of public accountability and good governance in Africa. The main perspective

adopted by the Delivering Service Indicators index is one of citizens accessing services and facing potential shortcomings in those services available to them. The index is thus presented as a Service Delivery Report Card on education and health. However, unlike traditional citizen report cards, it assembles objective information from micro level surveys of service delivery units.

---

This paper—a product of the Poverty Reduction and Economic Management, Africa Region; and Public Sector and Governance, Middle East and North Africa Region—is part of a larger effort in the departments to better understand governance issues in service delivery. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at [bgauthier@worldbank.org](mailto:bgauthier@worldbank.org), or [wwane@worldbank.org](mailto:wwane@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Delivering Service Indicators in Education and Health in Africa: A Proposal

Tessa Bold<sup>a</sup>, Bernard Gauthier<sup>b</sup>, Jakob Svensson<sup>c</sup> and Waly Wane<sup>d</sup>

**Acknowledgements:** This paper was prepared for the African Economic Research Consortium (AERC) Nairobi. We are grateful to Lida Bteddini, Sherri Hansen and Henri Savolainen for their excellent research assistance and to Olu Ajakaiye, Linda Frey, Mwangi Kimenyi, Ritva Reinikka and numerous World Bank Africa region, DECRG, PREM and HDN staff for insightful discussions and comments on previous drafts. The current version has also benefitted greatly from comments from referees.

---

<sup>a</sup> IIES, Stockholm University.

<sup>b</sup> World Bank and HEC Montréal

<sup>c</sup> IIES, Stockholm University and CEPR.

<sup>d</sup> World Bank.

## 1. INTRODUCTION

Access to quality services – in particular in health and education – has been recognized as fundamental for wellbeing and economic development (World Bank, 2003). However, in Africa and other developing countries, service delivery is often poor or nonexistent: schools and health clinics are not open when they are supposed to be; teachers and health workers are frequently absent from schools and clinics and, when present, spend a significant amount of time not serving the intended beneficiaries; equipment, even when available, is not used; drugs and vaccines are misused; and public funds are expropriated.<sup>1</sup> Unfortunately, poor people suffer the most when the public sector does not deliver.

Inadequate service delivery is also reflected in poor results in terms of health and education outcomes. In Africa, many health and education indicators are dismal. Given the rate of progress, current trends will not allow these countries to meet the 2015 Millennium Development Goals (MDG) targets (UNESCO, 2009).

To accelerate progress toward the MDGs, developing country governments, donors and NGOs have committed increasing resources to improve service delivery in social sectors. However, budget allocations alone are poor indicators of the actual quality of services in countries with weak institutions and systematic service delivery failures. Moreover, when the failures are systematic, there are no quick fixes and relying on governments to address them by themselves seems unrealistic, not least since many of these failures have largely been unaddressed for a long time. A complementary approach calls for empowering citizens and other actors to bring pressure on governments to reform the system. However, for this to work, citizens need to have access to hard evidence on service delivery performance. The Delivering Services Indicators is an attempt to start providing such information.

To date there is no robust, standardized data set to measure the actual quality of service delivery as experienced by the user in Africa. Existing indicators in social sectors are disparate and limited and focus is almost exclusively on development outcomes/outputs rather than on the service delivery systems that create those outcomes. In fact, no set of indicators exists for measuring constraints associated with the supply side of service delivery and the behavior of frontline providers, both of which have a direct impact on the quality of service citizens receive. Without consistent and accurate data on the quality of services, it is difficult for citizens or their governments to know which services are under-performing, consequently leaving little incentive for citizens and governments to act on.

The Delivering Service Indicators (DSI) seeks to provide a set of indices for benchmarking service delivery performance in education and health in Africa in order to track progress in and

---

<sup>1</sup> Chaudhury et al. (2006) and Reinikka and Svensson (2004), among others, provide systematic evidence of public service delivery failures. For additional evidence, see World Bank (2003).

across countries over time. It seeks to enhance effective and active monitoring and evaluation of service delivery and to become an instrument of public accountability and good governance in Africa. One of the key objectives of the DSI Index is to help reduce the gap of information between citizens, service providers and the state, which is at the root of poor service delivery performance, rent capture and misappropriation of resources. Ultimately, the purpose of the DSI Index is to help policymakers, citizens, service providers, donors and other stakeholders in enhancing the quality of service provision and ultimately improve development outcomes.

The perspective adopted by the DSI index is that of citizens accessing services and facing potential shortcomings in those services made available to them. The DSI index is thus presented as a Service Delivery Report Card on education and health. It seeks to measure the performance and quality of service delivery as experienced by citizens. However, instead of using citizens' perception to assess performance it assembles objective and quantitative information from a survey of service delivery units, using modules from Public Expenditure Tracking Survey (PETS), Quantitative Service Delivery Survey (QSDS), Staff Absenteeism Survey (SAS), and Observational studies.<sup>2</sup>

The index takes as its starting point the recent literature on how to boost education and health outcomes in developing countries. This literature shows robust evidence that incentives aimed at influencing the choice of effort exerted and the type of individuals attracted to specific tasks at different levels of the service delivery hierarchy, are positively and significantly related to education and health outcomes. In addition, conditional on providers exerting effort and being motivated, increasing resources can have beneficial effects. Therefore, we suggest that the indicators focus predominantly on measures that capture the outcome of efforts and type (broadly defined) both at the frontline service unit and by higher level authorities entrusted with the task of ensuring that schools and clinics are receiving proper support as well as the provision of physical resources to the service delivery unit. This choice also avoids the need to make strong structural assumptions on the link between inputs, behavior, and outcomes. While the data collection effort focuses on frontline providers, the indicators will mirror not only how the service delivery unit is performing but also indicate the efficacy of the whole institutional system in health and education. Importantly, we do not argue that we can measure these incentives directly, but, at best, measure the outcome of the set of incentives and constraints, including various supply constraints that influence performance using micro data.

---

<sup>2</sup> PETS trace the flow of public resources from the public budget to the intended final users through the administrative structure, as a means of ascertaining the extent to which the actual spending on services is consistent with the intended outcomes envisaged when budgets are allocated. QSDS examine activities and services at the provider level and the incentives and behavior of various agents in order to assess performance and efficiency of service delivery on the frontline. SAS focuses on the availability of teachers and health practitioners on the frontline in order to identify incentive problems and inefficiencies in resources utilization. An observational study aims at measuring the quality of services, proxied for by the level of effort exerted.

Over the past decade, new micro-level survey instruments (PETS, QSDS, SAS etc.), and dissemination tools like Citizen Report Cards, have proven to be powerful tools for identifying bottlenecks, inefficiencies and waste in service delivery, especially problems of resource leakage, delays, absenteeism, and inequities in the actual resource allocation. In the Ugandan education sector, for example, Reinikka and Svensson (2005, 2009) use a PETS approach to study the effects of a public information campaign aimed at empowering parents. They find a large reduction in resource leakage toward primary schools in response to the campaign. The introduction of a Citizen Report Card for Bangalore's public agencies led to major improvements in public satisfaction and the agencies' ratings. Similarly, corruption in the transactions between users and public officials declined markedly (Samuel, 2002). A randomized evaluation of a citizen report card intervention in the health sector in Uganda in 2005 concluded that the CRC led to significant improvements in the quality and quantity of primary health care provision and resulted in improved health outcomes in the communities (Björkman and Svensson, 2009).

This cumulative evidence provides grounds for hope that better information, particularly measurable and actionable indicators, will contribute to more accountability and transparency and ultimately to improved outcomes. The capacity to measure and compare key components of the main service delivery sectors over time and across countries should prove useful for various stakeholders. For citizens in particular, such information on public expenditure and service delivery performance would be especially valuable. As emphasized by CAFOD (2007), such information could help citizens: *“keep an active eye on government's progress and check whether policies are making a difference... help people give feedback to their government on the services they are providing... informed dialogue between government and citizens leads to more effective, fair and inclusive policies from which everyone benefits...”* (CAFOD, 2007, p.i)

This concept paper is structured as follows. Section 2 presents the citizen's perspective guiding the construction of the DSI Index and the main categories of indicators. It also presents the proposed indicators and the justification for their inclusion in the DSI index. Section 3 discusses the source of data to be used to construct the indicators. Section 4 presents the indicators' aggregation process and country ranking mechanisms and section 5 briefly discuss how the data could be presented and the pilot survey. Section 6 concludes.

## **2. DSI PERSPECTIVE**

In this section, we present the perspective guiding the construction of the DSI index. The categories of indicators are presented as well as the suggested indicators in each sector for measuring citizens' experience with service delivery.

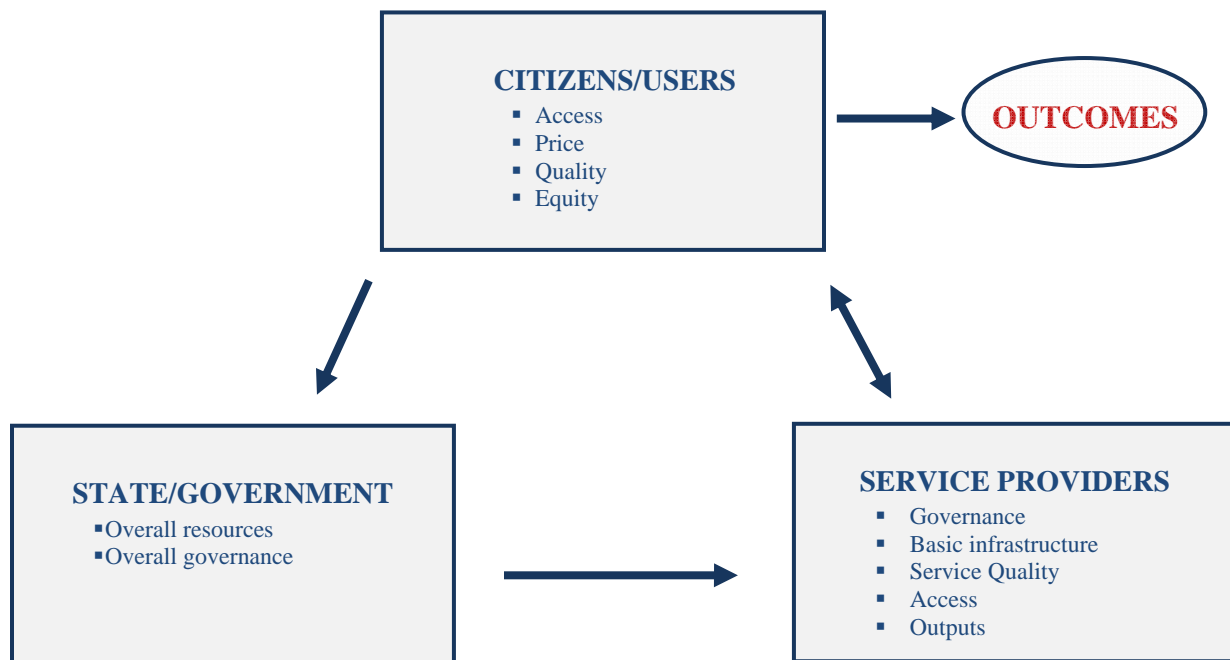
### **2.1 Service Delivery Outcomes and DSI Perspective**

Service delivery outcomes are determined by the interplay of government, providers and citizens. In particular, health and education outcomes are the results of the interaction between various actors involved in the multi-step service delivery systems and depend on characteristics and behavior of individuals and households. While delivery of quality health care and education is contingent foremost on what happens in consultation rooms and in classrooms, a combination of various basic elements have to be present in order for quality services to be accessible and produced by health personnel and teachers at the frontline, which all depend on the overall service delivery system and supply chain. Adequate financing, infrastructure, human resources, material and equipment have to be available for service delivery at the frontline, as well as proper institutions and governance structure to provide adequate incentives to the service providers. The availability of these essential elements and institutions are a function of the efficiency of the entire service delivery system.

Government at the upstream level sets overall policies, allocates resources and designs rules and service providers' incentive systems; service providers' behavior downstream is conditioned by these sets of constraints and incentives as well as rules determined locally; while citizens' decisions are influenced by the choices offered and the services' characteristics. Together, the choices and behavior of these various actors determine outcomes (Figure 1).

The DSI index seeks to measure key performance processes as experienced by the average citizen. Thus, the DSI index can be viewed as a Service Delivery Report Card on education and health. While emphasizing the experience of citizens at the service provider level, the index accounts for the various components of the supply side of service delivery and the multiple actors and multiple steps involved in service delivery systems. It builds on the Citizen Report Card approach but assembles information using one (composite) survey instrument, combining modules from PETS, QSDS, SAS, and observational surveys. Because health and education services are largely a government responsibility, and quite a lot of public resources have gone into these sectors, the DSI will at the initial stage focus on public providers.

**Figure 1: The Relationships between Citizens, Service Providers and Government**



## 2.2 DSI Categories and the Selection Criteria

There are a host of data sets available in both education and health. To a large extent, these data sets measure inputs and outcome/outputs in the service delivery process, mostly from a household perspective. While providing a wealth of information, existing data sets cover only a sub-sample of countries and are, in many cases, outdated (for instance there are five standard or interim DHS survey completed in Africa since 2007). Therefore, to compile an index of service delivery for Africa it is necessary to (i) have comparable data for all countries to be included in the index; and (ii) collect these data at the same regular intervals.

For the aforementioned reasons, it is difficult to rely on existing data sources (like DHS/LSMS/WMS). We therefore propose that all the data required for the indicator be collected through one standard instrument administered in all countries.

Our starting point is that service delivery can be measured using micro data. Moreover, while a ranking of countries on service delivery is obviously one of the key objectives of the DSI, country governments, and citizens, may be just as interested (or more) in benchmarking their own service providers; i.e., to make comparisons within a country. From the micro level data observations one can then build up inter-country comparisons.

Given this micro focus, we have essentially two options for collecting the indicators. We could

either take individuals or service providers as the unit of observation. We would argue that the only cost-effective option is to focus on service provider units. Obviously, this choice will to some extent restrict what type of data we can collect and indicators we can create.

More specifically, regarding the DSI index, we consider a service delivery production function,  $f$ , which maps physical inputs,  $x$ , the effort put in by the service provider  $e$ , as well as his/her type (or knowledge),  $\theta$ ,<sup>3</sup> to deliver quality services into individual level outcomes,  $y$ . The effort variable  $e$  could be thought of as a vector and thus include effort (broadly defined) of other actors in the service delivery system. Of course, as noted above, individual outcomes are not just affected by the service delivery unit, but also by the actions and behaviors of households,  $\varepsilon$ . We can therefore write

$$y = f(x, e, \theta) + \varepsilon . \quad (1)$$

To assess the quality of services provided one should therefore measure  $f(x, e, \theta)$ . Of course, the arguments in  $f$  are notoriously difficult to measure, and would involve a huge data collection effort. A more feasible approach is therefore to focus instead on the arguments which, to a first-order approximation, have the largest effects. The recent literature on the economics of education and health gives a fairly clear picture.

While there is little evidence that traditional input indicators (staff, material, infrastructure, etc.) are related to education and health outcomes in developed countries, there is some evidence that resource policies positively influence student performance in developing countries, hinting to diminishing returns for resources (Hanushek, 2003, p. 84). In a meta-analysis of 96 education studies during the 1980s in developing countries, several estimates indicate that increased infrastructure and material positively influence student performance (Hanushek, 2003, Table 6). More recent evidence from natural experiments and randomized evaluations indicate some potential positive effect of school resources on outcomes, but not uniformly positive (Duflo 2001; Glewwe and Kremer 2006).<sup>4</sup>

---

<sup>3</sup> We can think of type as the characteristic (knowledge) of the individuals self-selected into a specific tasks (or job).

<sup>4</sup> Case and Deaton (1999) show, using a natural experiment in South Africa, that increase in school resources (as measured by the student-teacher ratio) raises academic achievement among black students. They find that decreasing the student-teacher ratio by half raises students' reading test scores by an amount equivalent to the impact of two additional years of schooling (Hanushek, 2003). Also using a natural experiment, Duflo (2001) finds that a school construction policy in Indonesia was effective in increasing the quantity of education. She estimates that each school built for every 1,000 children led to an average increase of 0.12 years of education. Banerjee et al (2000) find, using a randomized evaluation in India, that provision of additional teachers in nonformal education centers increases school participation by girls. Furthermore, randomized trials in Nicaragua show that increase number of books has significant impacts on pupils' math scores (Jamison et al 1981). Similarly, Tan et al (1997) show that provision of textbooks raises performance on academic tests in the Philippines. However, a series of randomized evaluations in Kenya indicate that the only effect of textbooks on outcomes was among the better students (Glewwe and Kremer, 2006; Glewwe, Kremer and Moulin, 2002).

The somewhat weak relationship between resources and outcomes has been associated with deficiencies in the incentive structure of school and health systems. Indeed, most service delivery systems in developing countries present frontline providers with a set of incentives that negate the impact of pure resource-based policies. Therefore, while resources alone appear to have a somewhat limited impact on the quality of education (and health) in developing countries, coupling them with changes in incentives can have large and significant impacts on provider effort and the type and knowledge of staff at the frontline service units, which in turn positively affect outcomes. As noted by Duflo, Dupas, and Kremer (2009), the fact that budgets have not kept pace with enrollment, leading to large pupil-teacher ratios, over-stretched physical infrastructure, and insufficient number of textbooks, etc, is problematic, however increasing the level of resources might not help make up for the quality deficit in education and health without taking providers' incentives into account. In other words, the literature suggests that increasing provider effort and knowledge is essential to improve outcomes.

In sum, the literature points to the importance of provider behavior and knowledge in the delivery of health and education. However, conditional on service providers exerting effort and being motivated, it is clear that the provision of physical resources and infrastructure – especially in health – also has important effects on the quality of service delivery.

We therefore propose three sets of indicators : The first attempts to measure availability of key infrastructure and inputs at the frontline service provider level. The second attempts to measure effort and knowledge to deliver quality services at the frontline level. The third attempts to proxy for effort, broadly defined, higher up in the service delivery chain. Providing countries with detailed and comparable data on these important dimensions of service delivery is one of the main innovations of the DSI.<sup>5</sup>

In addition, we want indicators/measures that are (i) quantitative (to avoid problems of perception biases that limit both cross-country and over time comparisons)<sup>6</sup>, (ii) ordinal in nature (to allow within and cross-country comparisons); (iii) robust (in the sense that the methodology used to construct the indicators can be verified and replicated); (iv) actionable; and (v) cost effective.

---

<sup>5</sup> The suggested DSI set in education and health is based partially on an initial list of 50 PETS and QSDS indicators devised part of the project “Harmonization of Public Expenditure Tracking Surveys (PETS) and Quantitative Service delivery Surveys (QSDS) at the World Bank. This initial list, which covers a wide range of factors characterizing public expenditure and service delivery in the social sectors, was streamlined using the DSI criteria and conceptual framework (see Gauthier, 2008).

<sup>6</sup> See for instance Olken (2009).

## 2.3 Indicator Descriptions

Table 1 lists the indicators by sectors that have been identified.

**Table 1. A Service Delivery Report Card**

<b>Education</b>	<b>Health</b>
<i><b>At the school: Inputs and infrastructure</b></i>	<i><b>At the clinic: Inputs and infrastructure</b></i>
<ul style="list-style-type: none"> <li>a. infrastructure (electricity, water, sanitation)</li> <li>b. children per classroom</li> <li>c. student/teacher ratio</li> <li>d. books per student</li> </ul>	<ul style="list-style-type: none"> <li>a. infrastructure (electricity, water, sanitation)</li> <li>b. qualified health personnel per 1000 inhabitants</li> <li>c. medical material and equipment per clinic</li> </ul>
<i><b>Teachers: Effort and knowledge</b></i>	<i><b>Medical personnel: Effort and knowledge</b></i>
<ul style="list-style-type: none"> <li>e. Absence rate</li> <li>f. Number of hours children are being taught</li> <li>g. Share of teachers with minimum knowledge</li> </ul>	<ul style="list-style-type: none"> <li>d. Absence rate</li> <li>e. Time spent counseling patients per day</li> <li>f. Clinical performance in outpatient consultations</li> </ul>
<i><b>Funding: Effort in the supply chain</b></i>	<i><b>Funding: Effort in the supply chain</b></i>
<ul style="list-style-type: none"> <li>g. Effective education expenditure reaching primary school (in % of the primary education budget)</li> <li>h. Delays in wages</li> </ul>	<ul style="list-style-type: none"> <li>h. Effective primary health expenditure reaching primary clinics (in % of the primary health budget)</li> <li>i. Delays in wages</li> </ul>

The various indicators in each sector are briefly discussed below. A more detailed description and definition of the indicators are presented in a Technical Note available upon request (Bold, Gauthier, Svensson, and Wane, 2009).

## 2.4.1 Education

### At the school

#### *Infrastructure (electricity, water, sanitation)*

School buildings often lack basic infrastructure, in particular public schools in rural areas. The indicators account for the three basic infrastructures: availability of electricity, water and sanitation.

#### *Children per classroom*

In many countries or regions, students and teachers in schools lack access to basic primary school facilities and, in many cases, schools are overcrowded. The indicator of availability of classrooms is measured as the ratio of the number of primary school age children per available primary school classrooms.

#### *Students/teacher ratio*

Teacher shortage is a problem in many SSA countries and regions, especially in poorer and rural areas. UNESCO has estimated that about 145,000 teachers should be recruited – 77% above the observed increase between 1999 and 2006 -- to fill the needs of basic education annually to meet the MDGs. The indicator of teachers' availability is measured as average number of students per teacher.<sup>7</sup>

#### *Books per student*

Lack of basic education material is often an important constraint for learning faced by children and teachers in SSA countries. For instance, a SACMEQ II survey found that over half of grade 6 students in many SSA countries did not have a single book (e.g. Malawi, Mozambique, Uganda, and Zambia). Similarly, a PETS in Namibia reported only 0.58 mathematics textbooks per student (Schade and Naimhwaka, 2004, p. 67). The indicator of learning material is measured as the overall number of books available within primary schools per student.

### Teachers

#### *Absence rate*

In several low income countries, highly centralized personnel systems, inadequate incentives systems and weak local accountability have resulted in high levels of staff absenteeism. Recent

---

<sup>7</sup> As noted by UNESCO (2009), the student-teacher ratio could be viewed also as a proxy of education quality. The variation in learning outcomes In SACMEQ II for instance, it is explained at 36% by the student-teacher ratio.

findings on absence rates from a large multi-country study indicate that absenteeism rates among teachers are between 11 and 27 percent and translate into low quality of services (Chaudhury et al. 2006), and country-specific surveys indicate even larger rates (close to 50% in primary health clinics in Uganda, Bjorkman and Svensson, 2009). The indicator of absenteeism among the frontline teaching staff is measured as the share of teachers not in schools as observed during two unannounced visits.

#### *Number of hours children are in school being taught*

Even when in school, survey evidence from several countries suggests that teachers are not necessarily teaching (at regular teaching hours). The absenteeism survey can also be used to measure the extent to which teachers are not only in the school but actually in a classroom, as opposed to some other activity at the school. But plenty of anecdotal evidence suggests that even when a teacher is in a classroom, time is not necessarily spent on teaching. Through a three-pronged approach, we will capture the actual number of hours children are being taught. To this end we will combine data from the absenteeism survey, with information on actual (reported) teaching hours, and classroom observations to estimate the number of hours children are in school being taught (per day or week). The classroom observations will be done both in the classroom and by having the enumerators observe practice outside of it. As a by-product of the classroom observations we will generate detailed information about the pedagogical practices in the classroom.

#### *Share of teachers with minimum qualification*

Having teachers teaching, however, many not be enough if the teacher's competence (ability/knowledge) is inadequate -- a major problem in several SSA countries and especially in rural and poorer areas. This indicator measures teacher's knowledge and is based on a basic test of competence administered at the school level. The proposal is to test all (or at least a sample of 10) teachers teaching English (French) and mathematics in Grade 4, and those that have taught those subjects in Grade 3 the preceding year. This will be done by administering mathematics and/or language tests covering the primary curriculum as well as a pedagogy test adapted from teacher training entry tests in developed countries.

### Budget

#### *Effective education expenditure reaching primary schools (in % of the education budget)*

The transfer and availability of resources toward frontline service providers could be very low in various countries due to leakage of resources or because of allocation rules that favor administrative or other purposes not directly related to direct services to citizens. The indicator of availability of resources at the primary school level assesses the share of the education budget available for direct services to students at the school level. That is, for every dollar allocated in

the budget for a specific line ministry, the indicator measures the percentage that arrives at the provider level (and hence conversely, how much is used for other purposes, including capture and leakage, at other levels.). It is measured as the (non wage) recurrent expenditure reaching the frontline provider level in proportion of the sector's overall budget.<sup>8</sup>

#### *Delays in salaries*

Delays and bottlenecks could affect the allocation of resources through public administrations and potentially have important effects on the quality of services, staff morale and the capacity of providers to deliver services. The indicator of delays in salaries measures the proportion of teachers who have more than two months of salary overdue.

## **2.4.2 Health**

### At the clinic

#### *Infrastructure (electricity, water, sanitation)*

The access to basic infrastructure services to health facilities conditions the quality of services to the population. However, health clinics often lack the basic infrastructure, in particular in public clinics in rural areas. Electricity access, which is limited in several African countries, is important for various equipment usage, and overall use of facilities. Similarly, availability of clean water supply and improved sanitation at the facility level are fundamental for quality services given that un-cleaned water and inadequate elimination of used water are important vector of sickness. The indicators account for basic infrastructure; electricity, water and sanitation. It is constructed as the un-weighted average of availability of electricity, water and sanitation within primary health facilities.

#### *Medical material and equipment per clinic*

The lack of basic medical material and equipment is often an important constraint to accessibility and quality of health care services. The indicator of medical material is measured as the % of primary care providers where the following basic materials are available: thermometer, stethoscope, and weighting scale.

### Medical personnel

#### *Absence rate*

Similar as for education.

---

<sup>8</sup> Measuring effective education expenditure reaching primary schools is a challenging task since resource systems and flows differ a lot across countries, and receipts at the schools may be lacking.

### *Clinical performance in outpatient consultations*

The competence of health personnel is a major problem in several countries and especially in rural and poorer areas. Clinical performance in outpatient consultations will be measured through Patient Case Simulations (also called “health vignettes”). With this methodology one of the surveyors act as a case study patient with some specific symptoms. The clinician who is informed of the simulation is asked to proceed as if the enumerator is a real patient, while another enumerator acts as an observer. High quality performance in outpatient consultations entails at least the following: (i) To systematically arrive at a correct diagnosis (or preliminary diagnosis); (ii) To provide an appropriate treatment (or referral); and (iii) To reveal important information to the patient about which actions to take (e.g., how to take medicine, what to do if the patient does not get better, etc.). The methodology presents several advantages: (a) All clinicians are presented with the same case study patients, thus making it easier to compare performance across clinicians; (b) The method is quick to implement, and does not require waiting for patients with particular diagnoses; (c) We avoid intrusion and ethical issues that would arise if we were studying real patients’ cases. The method also has its drawbacks. The most important one is that the situation is a not a real one and that this may bias the results.<sup>9</sup>

### *Time spent counseling patients per day*

This measure will be based on aggregating data from the observational study of medical personnel. In the observational study the clinician is observed during a short period. Then combining data on number of patients treated per day (or week) with the observational data on the time spent on each patient, it is possible to calculate the total time spent counseling patients per day.

### Budget

#### *Effective health expenditure reaching primary clinics (in % of the health budget)*

Similar as for education.

#### *Delays in salaries*

Similar as for education.

### **2.4.3. Discussion**

Referring back to the service delivery production function (1), we are trying to measure effort

---

<sup>9</sup> However, several studies that have compared Patient Case Simulations with Direct Observation of clinical performance have concluded that there is a high correlation between the performance scores under these different alternatives (Leonard and Masatu, 2007, Das, Hammer and Leonard, 2008). On balance, we therefore think that this approach is suitable for measuring clinical performance.

and ability by frontline providers in health and education via the observation schedules and hours taught and via the teacher and medical tests. We propose to measure effective budget and delays in salaries as a proxy for the effort exerted by higher level authorities. Finally, physical resources are measured by infrastructure and material indicators. Obviously we do not observe effort or ability directly so our approach is an indirect way of getting at the key arguments in equation (1). Moreover, the indicators, by construction, do not speak to the important question of how to affect these arguments because we cannot directly measure the incentive constraints influencing staff behavior, and even if we could we would not be able to causally link them. In fact, the cause of the problem may not be closely connected to the appropriate solution, since the institutional system has, over time, influenced norms and thus behavior. For instance, it may well be the case that low pay, or inadequate supply of inputs, has eroded (non-pecuniary) motivation to exert high effort and created norms that do not reward high effort. Then increasing pay, or increasing supply, may just result in better paid and better supplied workers that still lack motivation to serve the public.

## **2.5 Outcomes**

To avoid making structural assumptions on the link between inputs, performance and outcomes, we do not suggest that outcomes should be part of the index. However, when the various sub-indicators and the final index are presented, it also makes sense to report, separately, on outcomes. In health there are measures for many countries at the national level, such as U5 mortality rates. Quantity outcomes in education are also available (various measures of flows and stock of schooling) for a large subset of countries. However, on quality there is no comparable (for many countries at least) data available. Thus, we suggest that student learning achievement is collected as part of the survey in education. This in itself is a great achievement.

Available evidence indicates that the level of learning tends to be very low in SSA countries. For instance, assessments of the reading capacity among grade 6 students in 12 SSA countries indicates that less than 25% of the children in 10 of the 12 countries tested reached the desirable level of reading literacy (SACMEQ, 2000-2002). Learning outcomes will be measured by student test scores in mathematics and language. We propose that younger cohorts are tested, partly because there is very little data on their achievement, partly because SACMEQ already tests students in higher grades, partly because the sample of kids in school becomes more and more self-selective as we go higher up when drop out rates are high, and partly because we know that cognitive ability is most malleable at younger ages (see Heckman and Cunha, 2007). Our proposal is thus to test a random sample of 10 children in grade 4 on a one-to-one basis. The test should consist of 10-20 questions per subject and the question difficulty should be calibrated to ensure adequate discrimination of the tests.

In terms of an indicator, the measure should be easily interpretable, for instance the share of children that can read a simple sentence and the share of children that can add up single digits.

## **2.6 Additional Data to Be Collected**

Apart from the core list of indicators, and test score data for students in education, we suggest that data on clinic and school characteristics are collected as part of the survey. These data could be collected at a very low marginal cost and would be useful in case one wants to add additional indicators based on the availability and quality of infrastructure and inputs, and for policy analysis and research using the data.

## **3. DATA SOURCE**

The data used to construct the DSI indicators will be based on facility surveys in health and education. All the necessary data needed to construct DSI indices would be collected through a single survey in each sector carried out at the frontline facility level, complemented by budget data collected at the central level.

The methodology to collect data on each of the indicators is based on international best practice for PETS, QSDS, SAS and observational surveys. The survey data come from direct interviews with school and health facility managers responsible for managing financial and material resources at the frontline and providing services, as well as frontline staff. Records from higher level government are also used to measure overall resource allocation. Such data collections are done independently from governments in order to ensure objectivity.

The data collected will be based on a representative sample of public primary providers. To ensure representativeness with a sample of 200 units, the sample will be stratified along relevant criteria (like urban-rural location).

## **4. CRITERIA FOR SAMPLE UNITS**

Primary care is delivered by different types units. In order to allow for a meaningful comparison across countries, it is important to define “similar” service delivery units in all countries. To this end we will use the following operational criteria:

- 1) The health unit provides primary health services defined as outpatient consultations, antenatal care, normal deliveries, immunization and family planning.
- 2) The health unit allocates at least 75% of its resources to primary health care services.
- 3) The health unit is the main source of primary health care for a significant proportion of the population (at least 25%).

In education the focus will be on government primary schools.

## 5. INDICATOR AGGREGATION PROCESS AND COUNTRY RANKING

The choice of a small set of easily interpretable indicators makes it possible to focus both on a direct comparison of the various sub-indicators as well as an aggregate score. On the latter, the DSI index is a summary measure of service delivery performance in education and health. In each sector, the DSI index measures the average achievement in a country across different categories. Each category is composed of two to four sub indicators measuring specific elements of performance within the sector.

The methodology used to construct the DSI index for allowing comparability and aggregation of indicators is as follows.<sup>10</sup>

Suppose there are  $k$  dimensions of inputs into quality service delivery in a sector. Let  $x_k^j$  denote the performance of country  $j$  along dimension  $k$ , and let the bundle  $X = (x_1, x_2, \dots, x_k) \in R^K$  summarize the performance across all dimensions (for a country).

The general indexing problem is about finding an appropriate indicator that maps  $X$  onto the real line, so that it can be ordered across countries. There are many approaches to aggregation, but they are all special cases of the following general class

$$I(X | \beta) = \left[ \frac{w_1 \Phi_1(x_1)^\beta + \dots + w_k \Phi_k(x_k)^\beta}{w_1 + \dots + w_k} \right]^{1/\beta} \quad (2)$$

The index is defined as the weighted mean of order  $\beta$  of the transformed outcomes  $\Phi_j(x_j)$ . The dimension weights  $w_j$  are all non-negative and can be assumed to sum to 1. Constructing the composite index therefore requires choosing the transformation functions,<sup>11</sup> the parameter  $\beta = 1 - 1/\sigma$ , where  $\sigma$  measures the elasticity of substitution between outcomes,<sup>12</sup> and an appropriate set of weights.<sup>13</sup>

---

<sup>10</sup> The discussion of how to construct composite welfare indices is based on a review article by Decanq and Lugo (2008).

<sup>11</sup> Transformation functions are used to make outcomes scale independent, and are also useful to deal with outliers and skewed measurements. Common transformation functions are: z-scores, distance to a reference point, indicators above or below the mean, evolution relative to trend, percentage annual difference over consecutive years or min-max.

<sup>12</sup> The smaller  $\beta$ , the smaller the allowed substitutability between dimensions and the lower the  $\beta$ , the more sensitive the indicator is to weak performing dimensions.

<sup>13</sup> The weights contain important information about the trade-offs between different dimension that society is willing to make.

Our preferred method is to use fixed weights for the aggregation purposes. Equal weights are the most commonly used approach for index construction.<sup>14</sup> While there is no analytical reason for this choice, equal weights are a common approach that is often justified on the basis of transparency.<sup>15</sup> We also suggest using the min-max transformation function  $\Phi(x_k) = [x_k - \min(x_k)] / [\max(x_k) - \min(x_k)]$  which transforms the variable into a uniformly distributed variable between 0 and 1.

For each category  $c$ , we can then create sub-indices  $X_1$  to  $X_3$ , where,

$$X_c = \frac{1}{K_c} \sum_1^{K_c} \Phi(x_k | category = c) \quad (3)$$

where  $K_c$  is number of dimensions in category  $c$ , and  $K = K_1 + K_2 + K_3$ .

Finally, we can take an equally weighted average of all the dimensions across the sub-indices to get at a DSI ranking within a sector; i.e.,

$$DSI = \frac{1}{K} \left[ \sum_1^{K_1} \Phi(x_k | category = 1) + \sum_1^{K_2} \Phi(x_k | category = 2) + \sum_1^{K_3} \Phi(x_k | category = 3) \right] \quad (4)$$

to get at the final index for a given sector.

Ultimately, there could be an overall aggregation of the three sectors index leading to an overall score and ranking of countries based on the three service delivery sectors.

In the aggregation process proposed above, we have implicitly assumed that all units sampled should carry the same weight. We could refine this further by weighing units by size of the catchment area. Thus, a clinic or school serving a smaller community would count less in the index than a clinic or school serving a larger community.

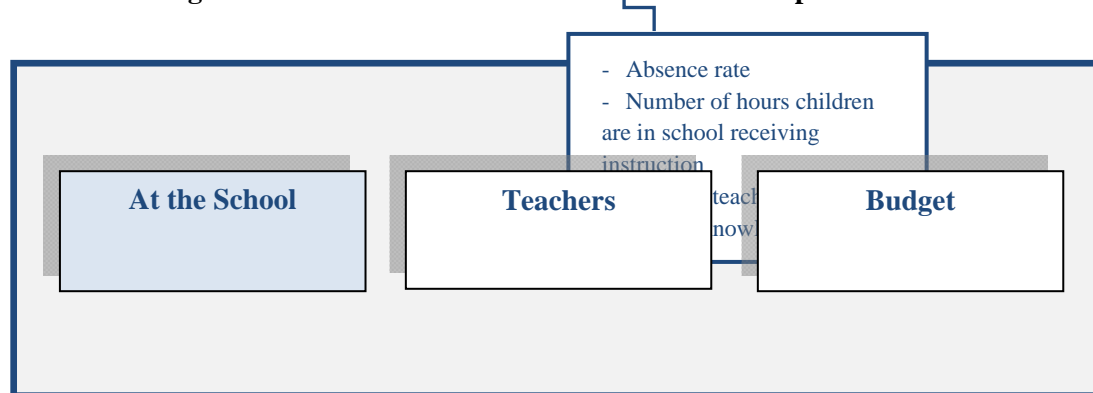
---

<sup>14</sup> See: UNDP Human Development Index, UNESCO Education for all index, World Bank Doing Business Index.

<sup>15</sup> Of course, they imply the assumption that there is a one-for-one trade-off between different dimensions. While equal weights are the most commonly used approach, there exist various data driven weighting schemes. One approach is to use frequency based weights, which are chosen relative to the proportion of the population suffering deprivation in a particular dimension. Alternatively, weights depend positively on the accuracy with which the data is measured. Another approach would be to use weights chosen to maximize the individual/country's performance in the index. This is sometimes viewed as appealing from a political point of view and can be interpreted as "giving everyone the benefit of the doubt". An alternative approach would be to use "normative weights", which are constructed on the basis of asking experts on how they would trade-off different dimensions using the budget allocation or analytic hierarchy process. Also, one could use various statistical techniques to estimate weights that summarize the data. Approaches based on principal component and factor analysis have been used to control for multicollinearity between different dimensions and produce weighting schemes that penalize highly correlated dimensions and give more weight to dimensions that are orthogonal. Overall, none of these approaches is perfect and any weighting scheme that arises from them should always be assessed according to the reasonableness of the implied trade-offs between dimensions (for further discussion, see for instance Decancq and Lugo, 2008).

Ease of presentation, access and interpretation of DSI indicators are important to favor their use by the various stakeholders. For each of the sectors, the information on the countries' percentile score and relative ranking for each of the sub-indices, aggregate categories and overall sector, would be provided on a DSI web platform. Figure 2 presents the DSI in education sector and their components as they would appear on a web platform.

**Figure 2: Delivering Service Indicators in Education and their Components**



*Illustration of the DSI in the Education Sector*

## 6. PILOT SURVEYS AND PRESENTATION

Following the design of the methodology for constructing the indicators, including survey instruments for collecting the data and implementation guidelines, AERC will pre-test the survey instruments in Tanzania and Senegal in the spring of 2010. Once the pre-tests are completed and the survey instruments refined, AERC will identify research teams and enumerators to pilot the first round of data collection in approximately 8 to 10 countries in Sub-Saharan Africa.

The pilot surveys should have two main objectives: (i) to field test and refine instruments that can be used to collect robust data; (ii) to identify a cost-effective and replicable strategy to collect data. After the first two pilots, the list of indicators could possibly be revised, particular in light of the financial implications of scaling up the project.

Ease of presentation, access and interpretation of DSI indicators are important to encourage their use by the various stakeholders. For each of the sectors, the information on the countries' percentile score and relative ranking in each of the sub-indices, aggregate categories and overall sector, would be provided on a DSI web platform. Data on actual outcomes should also be presented, as well as simple breakdowns (between urban-rural location for example) of the sample.

## **7. CONCLUSION AND NEXT STEPS**

The DSI seeks to provide a set of indices for benchmarking service delivery performance in two key social sectors in Africa, education and health, in order to track progress within and across countries over time. It seeks to enhance effective and active monitoring and evaluation of service delivery systems and to become an instrument of public accountability and good governance in Africa. The ultimate objective of the DSI index is to help policymakers, citizens, donors and other stakeholders in enhancing the quality of service provision and ultimately improve population outcomes.

The main perspective adopted by the DSI index is one of citizens accessing services and facing potential shortcomings in those services available to them. The DSI index is thus presented as a Service Delivery Report Card on education and health. While emphasizing the experience of an average citizen at the service provider level, the index accounts for the various components of the supply side of service delivery and the multi-actors and multi-steps involved in service delivery systems. It assembles information from composite micro level surveys in education and health.

The current version of the DSI encompasses only primary levels in the education and health sectors. In future versions of the DSI, it would be very useful to have such service delivery report cards developed for secondary and tertiary levels and encompass other key social sectors such as water and sanitation.

## References

- Amin, Samia and Nazmul Chaudhury (2008) “An Introduction to Methodologies for Measuring Service Delivery in Education” in Amin, Samia, Das Jishnu and Marcus Goldstein (editors) *Are you Being Served? New Tools for Measuring Service Delivery*, The World Bank, Washington, D.C.
- Arndt, Christiane and Carmen L. Romero (2008) “Review of the policy utility of the Worldwide Governance Indicators for the Central American Countries”, *Working paper 0108*, Working paper series on public sector management, The World Bank, Washington DC.
- Banerjee, Abhijit, Angus Deaton and Esther Duflo (2004), “Wealth, Health, and Health Service Delivery in Rural Rajasthan”, *American Economic Review Papers and Proceedings* 94 (2): 326–30.
- Banerjee, Abhijit, and Esther Duflo (2005), “Addressing Absence”, *Journal of Economic Perspectives* 20 (1): 117–32.
- Banerjee, Abhijit, Suraj Jacob, and Michael Kremer with Jenny Lanjouw and Peter Lanjouw (2000) “Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials,” mimeo, MIT.
- Banerjee, Sudeshna, Heather Skilling, Vivien Foster, Cecilia Briceño-Garmendia, Elvira Morella, and Tarik Chfadi (2008), “Africa Infrastructure Country Diagnostic: Ebbing Water, Surging Deficits: Urban Water Supply in Sub-Saharan Africa”, Background Paper 12, The World Bank, Washington D.C, June.
- Besley, Timothy and Maitreesh Ghatak (2006) “Reforming Service Delivery”, *Journal of African Economies* (16): 127-156.
- Bergeron, Gilles and Joy Miller Del Rosso (2001) “Food and Education Indicator Guide” *Indicator Guides Series*, Food and Nutrition Technical Assistance (FANTA), Academy for Educational Development, Washington, DC.
- Billig, P., Bendahmane, D and A. Swindale (1999) *Water and Sanitation Indicators Measurement Guide*, Indicator Guides Series Title 2, Food and Nutrition Technical Assistance, Academy for Educational Development, USAID, June
- Björkman, Martina, and Jakob Svensson (2009), “Power to the People: Evidence from a Randomized Field Experiment on Community-based Monitoring in Uganda”, *Quarterly Journal of Economics* 124 (2).
- Bold, T., B. Gauthier, J. Svensson, and W. Wane (2009), “Delivering Service Indicators in Education and Health in Africa: Technical Note, Unpublished Working Paper.

- CAFOD (2007) "Monitoring Government Policies: A Toolkit for Civil Society Organizations in Africa," Catholic Agency for Overseas Development, [www.cafod.org.uk](http://www.cafod.org.uk).
- Case, Anne and Angus Deaton (1999) "School Inputs and Educational Outcomes in South Africa," *Quarterly Journal of Economics*, 114(3): 1047-1085.
- Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan and Halsey Rogers (2006) "Missing in Action: Teacher and Health Worker Absence in Developing Countries", *Journal of Economic Perspectives*, 20 (1): 91-116.
- Cohen, Jessica Pascaline Dupas (2008), "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment", Poverty Action Lab, October.
- Das Gupta M, V. Gauri, and S. Khemani (2003), "Primary Health Care in Nigeria: Decentralized Service Delivery in the States of Lagos and Kogi" Africa Region Human Development Working Paper, Series No. 70, The World Bank, Washington D.C., September.
- Das, Jishnu, and Jeffrey Hammer, (2004) "Which Doctor? Combining Vignettes and Item-Response to Measure Doctor Quality," The World Bank, Washington D.C.
- Das J, Hammer J, and Leonard K (2008). "The Quality of Medical Advice in Low-Income Countries". *Journal of Economic Perspectives*, 22(2):93-114.
- Decancq K. and M.A. Lugo (2008) "Setting weights in multidimensional indices of well-being", *OPHI Working Paper No. 18*, August.
- Duflo, Esther (2001) "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4): 795-814.
- Duflo, Esther, Pascaline Dupas and Michael Kremer (2009) "Additional Resources versus Organizational Changes in Education: Experimental Evidence from Kenya", MIT, mimeo, May
- Filmer, Deon and Lant H. Pritchett (1999) "The Impact of Public Spending on Health: Does Money Matter?" *Social Science and Medicine*, 58: 247-258.
- Gauthier, Bernard (2008) "Harmonizing and Improving the Efficiency of PETS/QSDS", AFTKL, The World Bank, Washington, D.C. March, mimeo.
- Gauthier, Bernard and Ritva Reinikka (2008) "Methodological Approaches to the Study of Institutions and Service Delivery: A Review of PETS, QSDS and CRCS in Africa", African Economic Research Consortium (AERC) Framework paper.
- Gauthier, Bernard and Waly Wane (2009) "Leakage of Public Resources in the Health Sector: An Empirical Investigation of Chad", *Journal of African Economies* (18): 52-83

- Glewwe, Paul and Michael Kremer, (2006) "Schools, Teachers, and Education Outcomes in Developing Countries," in Hanushek, E and F. Welch (editors) *Handbook on the Economics of Education*, Chap 16, North Holland.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin (2002) "Textbooks and Test Scores: Evidence from a Randomized Evaluation in Kenya," Development Research Group, World Bank, Washington, DC.
- Gonzalez de Asis, Maria, Donald O'Leary, Per Ljung, and John Butterworth (2008), "Improving Transparency, Integrity, and Accountability in Water Supply and Sanitation: Action, Learning, and Experiences", The World Bank Institute and Transparency International, Washington D.C., June.
- Hanushek, Eric (2003) "The Failure of Input-Based Schooling Policies," *Economic Journal*, 113(February): F64-F98.
- Kaufmann D. and H. Kraay (2008) "Governance Indicators: Where Are we, Where we should Be Going?" *World Bank Research Observer* (23):1-30
- Khemani, Stuti (2006), "Can Information Campaigns Overcome Political Obstacles to Serving the Poor," World Bank, Development Research Group. Washington, D.C., mimeo.
- Leonard K., and M.C. Masatu (2007). "Variation In the Quality of Care Accessible to Rural Communities in Tanzania" *Health Affairs*, 26(2).
- (2008)."Moving from the lab to the field: Exploring scrutiny and duration effects in lab experiments" *Economics Letters*, 100(2):284-287.
- Maestad, Ottar, Gaute Torsvik and Arild Aakvik (2009) "Overworked? The Relationship Between Workload and Health Worker Performance in Rural Tanzania" Working paper in Economics no 02/09, Department of Economics, University of Bergen, March.
- Morella, Elvira, Vivien Foster, and Sudeshna Ghosh Banerjee, (2008) "Climbing the Ladder: The State of Sanitation in Sub-Saharan Africa," Africa Infrastructure Country Diagnostic The World Bank, Washington, D.C., June.
- OECD (2008) *Handbook on Constructing Composite Indicators: Methodology and User Guide*, Organization for Economic Co-operation and Development, Paris.
- (2009) *Measuring Government Activity*, Organization for Economic Co-operation and Development, JRC European Commission, Paris.
- Olken, Ben (2009) "Corruption Perceptions vs. Corruption Reality", *Journal of Public Economics* 93 (7-8): 950-964.

- Reid, Gary J. (2008) “Actionable Governance Indicators: Concept and Measurement” Administrative and Civil Service Reform (ACSR) Thematic Group, The World Bank, Washington DC, February, mimeo
- Reinikka, Ritva and Jakob Svensson (2004) “Local Capture: Evidence from a Central Government Transfer Program in Uganda”, *Quarterly Journal of Economics*, 119 (2): 1-28
- (2005), “Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda.” *Journal of the European Economic Association*, 3 (2-3): 259-267.
- (2006) “How corruption affects service delivery and what can be done about it” in Susan Rose Ackerman (ed) *International Handbook on the Economics of Corruption*, pp. 441-446, Edward Elgar Publishing, MA
- SACMEQ (2000-2002), Southern and Eastern Africa Consortium for Monitoring Educational Quality, [www.sacmeq.org](http://www.sacmeq.org)
- Samuel, Paul (2002), *Holding the State to Account: Citizen Monitoring in Action*, Bangalore: Books for Change.
- Schade, K. and E. Naimhwaka (2004), *Public Expenditure Tracking Survey (PETS) and Quantitative Service Delivery Survey (QSDS) – The Education Sector*, Windhoek: Office of the Auditor General
- Tan, Jee-Peng, Julia Lane, and Paul Coustere (1997) “Putting Inputs to Work in Elementary Schools: What Can Be Done in the Philippines?” *Economic Development and Cultural Change*, 45(4): 857-879.
- UNESCO (2009) *Education For All Global Monitoring Report 2009: Overcoming inequality: why governance matters*, UNESCO publishing and Oxford University Press.
- WHO (2006), *The African Regional Health Report 2006: The Health of the People*, The World Health Organization, Washington, D.C.
- WHO (2008), *UN Water Global Annual Assessment of Sanitation and Drinking Water*, Geneva
- WHO/UNICEF (2008) *Progress on Drinking Water and Sanitation: Special focus on sanitation*, Joint Monitoring Programme for Water Supply and Sanitation (JMP), UNICEF New York, WHO Geneva.
- World Bank (2003) *World Development Report 2004: Making Services Work for Poor People*, The World Bank and Oxford University Press, Washington, DC.
- World Bank (2006), *Getting Africa on Track to Meet the MDGs on Water and Sanitation: A Status Overview of Sixteen African Countries*, Water and Sanitation Program, December.

World Bank (2008) *Global Monitoring Report 2008, MDGs and the environment: agenda for inclusive and sustainable development*, The World Bank, Washington DC.

World Bank (2009), *World Development Indicators*, The World Bank, Washington D.C.