

Document of
The World Bank

PROGRAM DOCUMENT

FOR THE

PROGRAM AND TRUST FUND

SERVICE DELIVERY INDICATORS PROGRAM

November 7, 2011

This document is being made publicly available prior to approval. This does not imply a presumed outcome. This document may be updated following approval and the updated document will be made publicly available in accordance with the Bank's policy on Access to Information.

CURRENCY EQUIVALENTS

(Exchange Rate Effective September 1, 2011)

Norwegian Kroner (NOK) 5.42893	=	United States Dollar (USD) 1
	=	
British Pound (GBP) 0.60750	=	USD1
	=	
Special Drawing Rights (SDR) 0.61	=	USD1

FISCAL YEAR

July 1 – June 30

ABBREVIATIONS AND ACRONYMS

AERC	Africa Economic Research Consortium
AfDB	African Development Bank
ACET	Africa Center for Economic Transformation
AFTHD	Africa region Human Development Department
CBO	Community-based organization
CIH	Center for International Health, University of Bergen
CMI	Christian Michelsen Institute
CRES	<i>Centre de Recherche Economique et Sociale</i> , Senegal
DFID	Department for International Development
IDA	International Development Association
IIES	Institute for International Economic Studies, University of Stockholm
MDG	Millennium Development Goals
NGO	Non-government organization
NORAD	Norwegian Agency for Development Cooperation
REPOA	Research on Poverty Alleviation, Tanzania
SDI	Service Delivery Indicators

Vice President:	Obiageli K. Ezekwesili
Chief Economist:	Shantayanan Devarajan
Sector Director:	Ritva S. Reinikka
Task Team Leader:	Gayle H. Martin

Contents

Program Summary	v
I. Strategic Context	7
II. Program Development Objectives	11
Program Objectives	11
Target audience and beneficiaries	11
Results Framework	11
III. Program Description	12
Component 1: Collecting Benchmarking Indicators	12
Component 2: Outreach and Dissemination	13
Component 3: Capacity Building	15
Program Activities	15
IV. Implementation	16
Country Selection and Sequencing of Surveys	16
Data Collection and Survey Methodology	17
Service Delivery Indicators Trust Fund	18
V. Governance	19
Governance Structure	19
Arrangements for Monitoring	21
VI. Budgeting and Financing	22
Budget	22
Financing	23
ANNEX A. TECHNICAL ANNEX	24
Data Collection and Survey Methodology	24
Level of Aggregation and Sample Size	24
Repeated Cross-section versus Panel Survey	26
Frequency of Surveys	26
Data Analysis and Availability	27
Education Indicators	27
Health Service Indicators	31
Additional Data on Explanatory Factors	34
ANNEX B. ORGANIZATIONAL RESOURCES REQUIRED (PRELIMINARY)	36
ANNEX C. INFORMATION PERTAINING TO DATA ACCESS	38
ANNEX D. CONTRIBUTORS AND PROGRAM PREPARATION TEAM	40

Program Summary

Strategic Context. Africa faces daunting human development challenges. Despite large amounts of resources devoted to education and health many millennium development goals (MDGs) continue to lag. Most governments and development partners embrace the principle of evidence-based policy making, but the evidence is often lacking. In fact, there is little robust and representative evidence of what teachers and health workers do during a typical work-day, their levels of knowledge and skills, how teachers perform their teaching activities and how well health workers diagnose and treat their patients. While lots of data exist, they are often of poor quality, are not comparable across countries or over time because of lack of standardization, or are not the correct data (for example, to link inputs with outcomes). Importantly, no set of indicators is available for measuring service delivery and quality at schools and clinics from the citizens' perspective. Furthermore, what you don't measure you can't hold service providers accountable for. Without consistent and accurate information on the quality of services whose validity is trusted by all parties, it is difficult for citizens and politicians alike to assess how service providers are performing, to work towards corrective action, and ultimately bring about improvements in service delivery.

Objectives. The objectives cover three main areas: data collection, data dissemination and use, and capacity building. Specifically, the objectives of the Service Delivery Indicators (SDI) Program are to: (i) collect robust evidence on quality of education and health services over time and across countries; (ii) disseminate the data and create a high level of public awareness of the Service Delivery Indicators, nationally and internationally; and (iii) strengthen the capacity of institutions in Africa to conduct surveys and analyze the data generated. The longer term goal is to promote the use of the data by a wide variety of stakeholders toward the ultimate end of improving service delivery for human development.

Program Description. The vision is to implement the Service Indicator Surveys initially in 15–20 countries in Africa to be repeated with predictable frequency (about every 3 years). The Program focuses on a core set of indicators at two levels: (i) the knowledge and effort of service providers, i.e. what frontline service providers know and do; and (ii) the availability of key inputs at the frontline for effective service provision. The former is a key contribution of the Program, as no comprehensive data collection has been devoted to what is going on inside schools and health facilities. Taken together, the indicators provide a useful snapshot of actual performance as well as possible constraints that may undermine the delivery of quality services.

There are at least three ways in which the SDI Program can contribute towards the longer term goal of improving service delivery. First, robust evidence on quality of service provision will be an important input for countries to identify key constraints and push for experimentation with potential solutions. The data can also be used for monitoring and thus help countries track progress over time. Second, by making the information public, the service delivery indicators can help address the information asymmetry problem and be used as a tool to enhance public scrutiny and accountability. They provide clients with a benchmark that tells them where education and health services in their country or region fall short, and what they can expect their school or clinic to deliver. Third, by making information about service delivery performance public, the indicators could spur healthy competition to improve performance, both within and between countries.

An annual report on service delivery will be published that reports on the most recent service delivery data for the participating countries. The accompanying outreach and dissemination activities will create a high level of public awareness of the indicators and survey results. The results will be presented in different formats and through different channels to different audiences. To reach all target groups, the dissemination strategy at country level will include key types of outputs: (i) Each national survey will have a complete Program report including methodology, results and discussion of findings; (ii) Media reports and press releases and policy briefs targeting specific audiences; (iii) Presentations of indicator findings at key stakeholder forums, including high-level government meetings, relevant ministries, specific advocacy groups, research conferences, media, etc.

A Partnership Initiative. The Program started as a partnership initiative among the World Bank, the African Economic Research Consortium (AERC) and the William and Flora Hewlett Foundation, and subsequently the African Development Bank (AfDB) joined the partnership. Over time this partnership is expected to expand. Establishing an inclusive, yet nimble governance structure is key to securing broad buy-in of the various stakeholders, with efficient and effective financial and technical oversight.

Implementation Arrangements. The Bank will be the implementing agency for the first six-year phase of this ten-year Program. The Program will be implemented at the country-level by Country Implementing Organizations, in collaboration with a World Bank-based Program Management Team. The Program Management Team will be housed in the Africa region of the Bank, in particular in the Human Development department, AFTHD, under the leadership of the Sector Director. A Steering Committee with broad representation from various stakeholders will provide advice and guidance for the execution of the Program. The Steering Committee will have an Executive Committee that will comprise major donors (over a pre-determined threshold contribution level). The survey will be implemented in 15–20 countries. The country selection will be made with the aim to maximize the impact of the Program. High impact requires that the data are credible, and that there are *champions for change* within the countries that are—or may become—interested in using the data.

Role of the Bank. The Bank is well-suited for manage this effort for a variety of reasons: technical expertise and ability to exercise quality control; its access to governments and development partners; and its breadth of operations and physical presence in most African countries. The Bank has the ability to oversee fieldwork and analysis by independent national institutions, and its fiduciary systems are valued by development partners who use Bank trust funds as vehicles to channel resources to priority development initiatives. A further rationale for Bank involvement is related to the institution’s role as a knowledge bank, and in this regard the partnership with the AfDB and the AERC amplifies this role.

I. Strategic Context¹

1. **Human development indicators continue to lag even where income poverty has been reduced.** The most recent assessments of progress toward the Millennium Development Goals (MDG) showed significant improvements in poverty reduction, but this is not always the case for the education- and health-related MDGs. While economic improvements are important for human development, they are not always sufficient.² This is particularly true for MDGs that are critically linked to the performance of service providers, and the underlying systems that determine or influence their capacity and behavior.

2. **The quality agenda—*How well are you being served?*** Africa faces major human development challenges. The access, use, as well as quality of health services are at worryingly low levels. In education, while enrollment has expanded very rapidly in many countries, quality is clearly lagging behind. To accelerate progress, developing country governments, donors, and non-governmental organizations (NGOs) have committed increasing resources to improve service delivery.³ In order to ensure that the returns to these and other human development investments are fully realized, it is not enough to ask “Are you being served?”—the title of a landmark publication on service delivery—but “How *well* are you being served?”⁴

3. **Accountability for public resources.** Developing country governments allocate roughly a third of their budgets to education and health. Demands for accountability and for the efficient use of public resources—from tax payers in developed or developing countries alike—are gaining in prominence. The global economic situation has forced a reexamination of public spending and the demand for accountability is likely to increase in the coming years.

4. **What gets measured gets managed.** Most governments and development partners embrace the principle of evidence-based policy making, but detailed evidence of the obstacles to improvement is often missing. In fact, there is little robust and representative evidence of what teachers and health workers do during a typical work-day, their levels of knowledge and skills, how teachers perform their teaching activities and how well health workers diagnose and treat their patients. Furthermore, what you don’t measure you can’t hold providers of services accountable for. Without consistent and accurate information on the quality of services whose validity is trusted by all parties, it is difficult for citizens or politicians to assess how service providers are performing and to work towards corrective action, and ultimately bring about improvements in service delivery.

¹ This document is largely based on the document prepared by the Chr. Michelson Institute (CMI), with generous funding from the Hewlett Foundation.

² World Bank. 2011. *Global Monitoring Report 2011: Improving the Odds of Achieving the MDGs*. World Bank, Washington, DC.

³ For example, since the 1990s, development assistance for health has quadrupled reaching USD27 billion in 2009. Health spending by developing country governments has also peaked, with spending nearly doubling to reach USD240 million in 2006. Similarly, development assistance for education doubled since 2002 reaching a high of USD10.8 billion in 2007 (World Bank, 2011).

⁴ Fiszbein, A., D. Ringold, and H.F. Rogers. 2011. “Making Services Work: Indicators, Assessments, and Benchmarking of Quality and Governance of Public Service Delivery in the Human Development Sectors.” *Policy Research Working Paper 5690*. World Bank, Washington DC.

5. **Translating the research on service delivery into operational gains and improved outcomes.** The 2004 World Development Report was visionary in focusing attention on front-line service providers and the relationships of accountability between clients, policymakers and providers.⁵ It pulled together technical work on service delivery and spurred further work on various themes: governance, accountability, quality of service provision, provider behavior and incentives, consumer behavior, voice, exit, etc. Given the transaction-intensive and discretionary nature of service delivery in education and health, the lack of systematic inclusion of these advances in the human development sectors is a glaring omission in the race toward the MDGs.

6. **What is wrong with the data we have?** Some of the weaknesses in the data that are currently available are:

- Lack of data linking inputs with outcomes. The focus of many data systems is on either on final outcomes or inputs, and not on measures that the service providers can be held accountable for.
- Lack of comparable data to assess performance across national or sub-national boundaries or over time. This result is considerable inefficiency—lots of data are collected, but data collection resources are wasted because of lack of standardization.
- Service delivery data are often of weak quality. Education and health sectors routinely collect administrative data (e.g. from education and health management information systems), but the data quality is highly variable, as is coverage.
- Administrative data are a potentially rich source of data, but have seldom delivered the quality and coverage of service delivery units needed to credibly track performance. A manifestation of the quality weaknesses of administrative data is that global reports (which are often based on aggregation of administrative data) and the situation on the ground (where a survey is done) often diverge.⁶

Importantly, no set of indicators is available for measuring constraints associated with service quality at the frontline (at schools and clinics) from the citizens' perspective.

7. **Making human development services work.** The above factors led to the conceptualization of an Africa-wide initiative that benchmarks human development service delivery performance, the Service Delivery Indicators (SDI) Program. It is envisaged that the broad availability of a robust set of service delivery indicators will heighten public awareness, help address the asymmetry of information, and form a basis for informed interaction among citizens, community-based organizations (CBOs), NGOs, policymakers, and service providers to improve the quality of services and ultimately to improve development outcomes. The indicators will also enable governments and service providers to identify gaps and monitor progress over time—a key to successful reforms.

8. **Institutionalizing benchmarking.** Recognizing the data gaps, the SDI Program aims to institutionalize the collection of a set of indicators that would gauge the quality of service delivery over time and across countries in Africa. The proposed set of indicators builds on a

⁵ World Bank. 2003. *World Development Report 2004: Making Services Work for Poor People*. World Bank. Washington, DC.

⁶ A study examined the number of children receiving routine immunizations in 193 countries from 1986 to 2006. The study revealed troubling gaps between the immunization data reported by governments and numbers based on independent surveys. Lim, S. D. Stein, A. Charrow, C. Murray. 2008. "Tracking progress towards universal childhood immunization and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunization coverage." *Lancet* 372 (2031-2046).

growing body of literature on measuring the performance of teachers and health providers. The choice of indicators takes as its starting point the importance of service provider competence and behavior. As long as teachers and health workers have a minimum level of skill and exert the necessary effort, there is evidence that the provision of other inputs and infrastructure has important effects on learning and quality of health services. Simply increasing the level of resources will not improve outcomes without also taking provider behavior into account. For this reason, the proposed metrics aim at measuring knowledge and effort of teachers and health workers at the frontline level. The selected indicators share four characteristics. They are: (i) quantitative (to avoid problems of perception biases that limit both cross-country and longitudinal comparisons); (ii) ordinal (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); and (iv) actionable.

9. **Finding pockets of innovation.** The service delivery information is likely to point to important gaps and failing performance. More importantly, it will also highlight positive outliers—examples of good performance in spite of financial and human resource constraints. They will provide an important source of learning because such examples will point to what is possible within the a country’s institutional and resource constraints. In doing so, the Program will enrich the somewhat simplistic common recommendation that more resources are needed to address the education and health system weaknesses. It is often not just resources that are needed.

10. **Rationale for Bank involvement.** As mentioned, significant resources that are invested in education and health in Africa by the countries themselves and by development partners. The SDI Program will generate important information that can help understand why there is a great variation in the progress toward the MDGs. Equally important is the institution’s role as a knowledge bank. It is central to the Bank’s mandate to generate information, contribute to learning with the ultimate objective of improving development effectiveness. In this regard the mandates of the World Bank and AfDB are similar, and the partnership with the AfDB in this initiative is therefore mutually reinforcing.

11. The World Bank will be the implementing agency for the first half of this ten-year Program. The Bank is well-suited for managing this effort for a variety of reasons: technical expertise and ability to exercise quality control; its access to governments and development partners; and its breadth of operation and presence in most African countries. The Bank has the ability to oversee fieldwork and analysis by independent national institutions, and its fiduciary systems are valued by development partners who use Bank trust funds as vehicles to channel resources to priority development initiatives. The Bank has recently taken stock of trust fund experiences, and efforts to improve efficiency and effectiveness of trust fund operations will be incorporated into this Program.⁷ Furthermore, the design of this trust fund has some unique design aspects that mitigate many of the risks raised in the trust fund stocktaking (discussed later).

⁷World Bank. 2011. *2010 Trust Fund Annual Report*. Global Partnership and Trust Fund Operations Department, Concessional Finance and Global Partnerships. World Bank, Washington, DC.
Independent Evaluation Group 2011. *Trust Fund Support for Development: An Evaluation of the World Bank’s Trust Fund Portfolio*. World Bank, Washington, DC.

12. **Link with key Bank strategic documents.** The SDI Program is a bold initiative that (i) offers a systematic approach to tracking service delivery performance and publicizing it, and to (ii) to work with governments and development partners to use the information and its operations to strategically and innovatively to improve the performance of services for the poor especially, to the end of improving human development outcomes. This is fully aligned with the Bank's mandate and its development priorities. The Program is furthermore consistent with the vision for building human capital in the Bank's Africa Strategy, *Africa's Future and the World Bank's Support to It?*⁸

13. **Relation to other surveys and data initiatives.** There exists a host of data sets on education and health. Many of the data sets, like Demographic and Health Surveys (DHS), Living Standards Measurement Studies (LSMS), Welfare Monitoring Surveys (WMS), and the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ), focus primarily on outcomes. Other data sets, such as Service Availability Mappings (SAM), focus only on inputs. These sources are often incomplete in their national coverage, the indicators are inconsistently defined and not comparable across countries, and the data quality is variable. In the case of facility surveys there are numerous examples, especially in the health sector, that follow different methodologies given their different purpose. In addition, there are routine data collection efforts – of varying levels of quality – that provide data on outputs. Importantly, none of these data sets provide any direct measure of what frontline service providers know and of the quality of their performance. The Service Delivery Indicators Program is a primary data collection initiative that fills this crucial gap. It is not the intention to replace the administrative data collection, but to be complement it, specifically none of these data sets provides a direct measure of the knowledge or behavior of frontline service providers that is captured consistently over time and across countries. And since the Indicators are based on a single, standard survey instrument that will be used in all countries, and since the costs of data collection are relatively low, the Indicators can be produced more frequently and with greater cross-national comparability than most other data sets.

14. **Relation to results-based financing.** There is currently increased interest in the use of Results-based Financing in development assistance. For instance, the Governments of Norway and United Kingdom have committed over USD550 million to results-based financing in the health sector through the World Bank's Health Results Innovation Trust Fund. One potential challenge of this approach is that since financing usually is not tied to outcomes, but rather to some intermediary output, it is not always clear how the financial incentives will affect the quality of the service. The focus of the Service Delivery Indicators Program on the quality of the service that is delivered provides a timely and useful complement to these innovative financing mechanisms.

⁸ World Bank. 2011. *Africa's Future and the World Bank's Support to It?* World Bank, Washington, DC. (p28).

II. Program Development Objectives

Program Objectives

15. The objectives cover three main areas: data collection, data dissemination and use, and capacity building. Specifically, the objectives of the SDI Program, are to: (i) collect robust evidence on quality of education and health services over time and across countries; (ii) disseminate the data and create a high level of public awareness of the Service Delivery Indicators, nationally and internationally; and (iii) strengthen the capacity of institutions in Africa to conduct surveys and analyze the data generated. The goal is to promote the use of the data by a wide variety of stakeholders toward the ultimate end of improving the quality of education and health services.

16. There are at least three ways in which the SDI Program can contribute towards the longer term goal of improving service delivery. First, robust evidence on quality of service provision will be an important input for countries in order to identify key constraints and push for experimentation with solutions. The data can also be used for monitoring and thus help countries track progress over time. Second, by making the information public, the service delivery indicators can help address the information asymmetry problem and be used as a tool to enhance public scrutiny and accountability. They provide clients with a benchmark that tells them where education and health services in their country or region fall short, and what they can expect their schools or clinics to deliver. Third, by making information about service delivery performance public, the indicators could spur healthy competition to improve performance, both within and between countries.

Target audience and beneficiaries

17. The indicators are selected in a way that makes them easily comprehensible and informative for citizens, civil society and advocacy groups, and the media. Furthermore, the indicators generated through the Program, as well as the more detailed, underlying data, speak to all levels of the service delivery chain, especially the government and implementing organizations from the ministerial level to the local government and to frontline service providers. Finally, the Program will inform international agencies, including donors, as well as training institutions and professional organizations, all of which are important actors within the service delivery chain.

Results Framework

18. In support of the objective of SDI, intermediate outcome indicators have been identified:

- (i) Public debate on education and health service delivery initiated and/or informed.
- (ii) Stakeholders (policymakers, media, NGOs, CSO) reporting use of SDI analysis within 6 months after any of the SDI dissemination events.

19. Ultimately, the Program outcome indicators are expected to contribute to the higher objective of improved service delivery. The basic premise is that reliable information can motivate for change. In many countries, both governments and citizens are aware that there are

quality deficiencies in the provision of education and health services. However, the lack of reliable, quantifiable measurement of the magnitude and the sources of the problems prevent an informed public discussion on the issues. There is power in information: with the regular publication of a credible set of service delivery indicators, citizens will know that government officials have this information, and likewise governments will know that citizens have this information. When data are available, it will be much more difficult to ignore serious quality problems. Equally important, these surveys will also provide concrete evidence where there are improvements in service quality.

20. The Service Delivery Indicators are likely to inspire action at various levels, depending on the country context. In some places, widespread media coverage and citizen mobilization may pressure recalcitrant government officials to act. In others, well-intentioned civil servants may push for change from within with measures to improve management and supervision, and enforce existing rules and regulations. Or in countries where high-level elected leaders (including the President) are politically motivated to demonstrate improvements in basic service delivery, the publicity associated with the Service Delivery Indicators may well motivate specific actions from the very highest levels of government. International policy priorities, including donor support, can also be altered towards supporting more effective measures for improved service quality. Finally, the indicators may provide important input to professional associations, which define and maintain professional standards, and for training institutions, which play a key role in shaping future frontline service providers.

III. Program Description

Component 1: Collecting Benchmarking Indicators

21. **Core Service Delivery Indicators.** As mentioned, the delivery of quality education and health services depends first and foremost on what happens in health centers and in classrooms. The ability of frontline service providers to deliver quality services depends in turn on efforts at higher levels of the supply chain to bring about the required resources (both in terms of knowledge, staff, equipment and other infrastructure), as well as in holding actors at all levels of the supply chain accountable for meeting performance targets.

Table 1. Core Indicators

<i>Knowledge and Effort of Service Providers</i>	
<i>Teachers</i>	<i>Health workers</i>
E1: Absence from school	H1: Absence from the health facility
E2: Absence from classroom	H2: Skills to reach correct diagnosis
E3: Share of teachers with minimum knowledge	H3: Skills to handle life-threatening complications for newborns and mothers
E4: Time spent teaching in the classroom	H4: Adherence to clinical guidelines
E5: Quality of instructions	
<i>Availability of Key Inputs</i>	
<i>Schools</i>	<i>Health facilities</i>
E6: Minimum teaching equipment available	H5: Availability of medical equipment
E7: Student-teacher ratio	H6: Drugs in stock
E8: Textbooks per student	H7: Workload per clinician
E9: School infrastructure	H8: Health facility infrastructure

22. Therefore the Program focuses on indicators at two levels: (i) the knowledge and effort of service providers, i.e. what frontline service providers know and do; and (ii) the availability of key inputs at the frontline for effective service provision. The former is a key contribution of the Program, as no comprehensive data collection has been devoted to what is going on inside service provider institutions. Taken together, the indicators provide a useful snapshot of actual performance as well as possible constraints that may undermine the delivery of quality services.

23. The choice of indicators is central to the Program. The process for selecting the indicators was extensive, detailed and widely consultative involving leading technical experts in the relevant fields. The Service Delivery Indicators were piloted in Tanzania and Senegal during 2010.⁹ The pilot demonstrated that the methodology is capable, through a single set of instruments and at a single point of collection, to provide information to construct a set of indices for benchmarking service delivery performance in education and health in Africa. Furthermore, a technical meeting involving global experts was held in November 2010 and the support for the technical merit as well as the importance of the initiative was confirmed. This was followed by a further process of stakeholder consultation and revision, and at a technical meeting (July 2011) convened by the Chr. Michelson Institute the indicators were presented and adopted.

24. Some modifications to the indicators are being piloted before scaling up the Program. Further consultations with sector specialists are being conducted order to ensure alignment and true complementarity with existing initiatives. The pilots also yielded important lessons for implementation and other practical aspects that have informed the design of the larger Program.

25. **Data access.** The Program is committed to making the results available as well as the data in disaggregated form as early as possible to allow use and analysis of the data by all stakeholders. This principle is consistent with the Bank's Open Data Policy. Detailed information is provided of the Bank's Open data Policy in ANNEX C. Under this policy, users are free to copy, distribute, adapt, display or include the data in other products for commercial and noncommercial purposes at no cost subject to certain limitations. The Program is also aware of the ethical concerns about protecting individuals participating in the survey, and there may be restrictions on making available personal or some geographic identifiers in datasets.¹⁰ This is common practice in other data surveys (e.g. Demographic and Health Surveys etc.). The World Bank's Microdata Library is one possible location where the data can be housed. It operates as a portal and disseminates micro-datasets from two sources: those that belong to the World Bank and those where the data have been generated and are owned by another agency.

Component 2: Outreach and Dissemination

⁹ It was decided that the pilots should include an Anglophone and Francophone country with different budget systems. The selection of Senegal and Tanzania was also influenced by the presence of strong local research institutes from the AERC network: *Centre de Recherche Economique et Sociale* (CRES) in Senegal and the Research on Poverty Alleviation (REPOA) in Tanzania. Both research institutes have extensive facility survey experience and are also grantees of the Hewlett-Packard supported Think Tank Initiative.

¹⁰ Unlike household surveys, we will be surveying clinics and schools per country (about 200-300 of each facility type). There may be some instances where a geographic identifier (say district name) may identify a facility if there is only one school or clinic per district, and only one provider (head teacher or doctor) per facility. This may lead to identification of a specific provider.

26. Data collection will be preceded by extensive consultation, and create a shared understanding for the rationale and trust in the quality of the data through dialogue and through transparent relations between the Program and the target groups. This consultation will facilitate more effective dissemination and uptake of the results.

27. **Country-level.** The aim of the outreach and dissemination activities is to create high level of public awareness of the core indicators. The results should be presented in different formats and through different channels to different audiences. To reach all target groups, the dissemination strategy at country level will include key types of outputs:

- Each national survey will have a complete Program report including methodology, results and discussion of findings. These reports are the basis for all other products developed.
- Media reports and press releases and policy briefs targeting specific audiences.
- Presentations of indicator findings at key stakeholder forums, including high-level government meetings, relevant ministries, specific advocacy groups, research conferences, etc. These presentations will focus on fostering dialogue around the findings and their implications. Presentations and discussions of the results should also be held in regions or districts where data have been collected.

For the national dissemination, close cooperation will be sought with international agencies and organizations that effectively can champion the dissemination efforts. These organizations need to be able to reach out effectively to a wide range of the identified target groups.

28. **International level.** The main dissemination output at the international level will be:

- A website where country results are presented and visualized in a user-friendly format. The website can also be expanded to an interactive forum where service providers and beneficiaries can voice their own experiences with service quality. The website will have links to the World Bank Microdata website,¹¹ and where all data are available for further analysis.
- An annual Service Delivery Indicators report. This status report will progressively have data for all countries as data collection expands. Not all countries will have a survey in a given year. Therefore, the annual report will present the most current data of the respective service delivery core indicators across all countries involved in the Program.
- A detailed Communication Strategy will accompany each annual report, including presentations at global and regional policy forums, media reports and press releases and policy briefs targeting specific audiences.

29. A key principal is that all data will be publicly available with the speediest turnaround time. This will increase trust and will facilitate wider use of the data by research institutions and think-thanks.

30. **Interacting with the Providers and Beneficiaries.** It is envisaged that data collection gradually will move towards electronic platforms. This will enable real-time feedback to service providers about their level of performance. It is also possible to set up national servers that can be called upon during field visits, enabling each health facility and school to assess their relative

¹¹ <http://microdata.worldbank.org/>

performance compared to other service providers. Such direct feedback to the service providers could be an important mechanism to motivate for change.

31. A further possible extension would be to incentivize service providers (for instance through free airtime for mobile phones) to respond to small surveys on key aspects of service quality via the mobile networks. Such data collection is clearly less reliable than information collected by external teams, but the very process of data collection may be an important mechanism for retaining focus on service quality at the frontline. Adding an interactive part to the Program is also envisaged to create increased interest and ownership.

32. Finally, the Program may engage with beneficiaries of the services through the SDI website, mobile phones, Facebook and other social media to voice their experience and satisfaction with education and health services. One may for instance put up a short questionnaire or ask for comments to the SDI results. This information can itself become useful for advocacy purposes, but more importantly, such a bottom-up approach is likely to increase the legitimacy and interest in the Program among civil society, media and policy stakeholders.

Component 3: Capacity Building.

33. Capacity building is one of the main expected outcomes of the SDI Program. The capacity of national research and policy institutions to design and implement surveys will be strengthened through their role as implementing partners for the SDI surveys. They will work in close collaboration with technical experts from the Bank, external institutions and national specialists in managing the surveys as was done in the pilots with think tanks such as Research on Poverty Alleviation (REPOA) in Tanzania and the *Centre de Recherche Economique et Sociale* (CRES) in Senegal. The technical aspects of the surveys will be modified in each country in collaboration with the national partners giving them access and hands-on experience of managing survey methodology and tools. They will participate in producing knowledge products and have quick access to the data generated. The capacity of national partners will also be strengthened through their leading role in disseminating the results of the survey and conducting dialogue and advocacy sessions with national stakeholders and decision makers.

34. To ensure that capacity building is realized, the terms of reference of the country advisers will reflect specific skills and tasks for this role. In addition, the performance indicators of the SDI Program require a national expert to be the co-principal investigator of the surveys. National partners will also be exposed to the survey process and outcomes of other participating countries for learning purposes.

Program Activities

35. **Error! Reference source not found.** lists some of the activities to be financed. These include:

- Refinement of the Service Delivery Indicators and related preparation for start-up of the Program.
- Implementation of the indicator surveys and data collection.

- Preparation of a high profile report that benchmarks education and health service delivery in Africa (formal title to be finalized) that will be distributed annually.
- Various knowledge products will be prepared. Examples include products listed in the Table below. A key target audience for these products will be AERC affiliates and educational institutions in survey countries.
- Trust Fund Administration and Management.

Table 2. Possible Knowledge Products

Knowledge Products	Content Definitions
Background Briefs	Background and concept information about the SDI Program, partnerships and the data.
Case Studies	Case studies of service delivery performance and innovations. Will include country and sectoral context service delivery performance, service delivery challenges, new innovations in service delivery (including design, implementation and evaluation).
Country Examples	Short synopses of the main features of service delivery benchmarking and innovations.
Feature articles, presentations and technical briefs, technical working papers	These products will put to use the service delivery indicators, and the product will vary depending on the audience.
Tools and Guidelines	Working aids for health workers on the ground e.g. How-to tools for designing and implementing service delivery surveys and guidelines for analysis.

IV. Implementation

36. An Implementation Manual will developed that provides more detail about the implementation plans, implementation arrangements, detailed description of the terms of reference for the various positions and structures in the organization design, the specific survey instruments etc. The plan will also include the specific activities, timelines, as well as a procurement plan.

Country Selection and Sequencing of Surveys

37. The survey will be implemented in 15–20 countries. The country selection will be made with the aim to maximize the impact of the Program. High impact requires first and foremost that the data are credible, and secondly that there are “champions of change” within the countries that are—or may become—interested in using the data. Furthermore, in order to facilitate healthy competition between countries, each country must be able to compare itself with other countries that are seen as relevant comparisons.

38. Some of the country selection criteria that will be considered are:

- The existence of local institutions capable of implementing the survey with sufficiently high quality.
- Contextual factors that influence the likely impact of the Program at the country level.
- The significance of the country as a relevant comparison to other countries (taking into account factors such as country size, level of development, political stability, governance structure, post-conflict situation, etc.).

- Geographical location (East, West, Central and South) to ensure the pan-African vision of the Program.
- Main language (first implement the Program in Anglophone and Francophone countries to capitalize on existing tools developed during the pilot phase and subsequently add Lusophone countries).

39. Assuming that a new survey is produced in each country every third year, there will be three survey waves, each wave encompassing one third of the countries in the Program. As a point of departure, we suggest five countries in each wave, i.e., 15 countries in total. The number of countries can be expanded as additional resources become available.

Data Collection and Survey Methodology

40. All data will be collected at the health facility or school level. In the education sector, focus will be on primary schools, while in the health sector to focus will be on primary health services up to the level of first referral hospitals (outpatient services and maternal and newborn services). Government, non-profit and private for-profit facilities will be included as appropriate.¹² The data will be gathered from direct observation of provider behavior, from various test of provider knowledge and skills, and from observation of the availability of key inputs required to enable provision of quality services (infrastructure, equipment, supplies, etc.).

41. The pilot developed detailed manuals and data collection tools for the indicators. Any new indicators, and those that have been revised during the consultation process, need to be piloted before implementation at scale. Data will be collected by trained enumerators, one team for each sector (education and health supervised by a survey leader, again under the supervision of the country leader within the implementing organization in each country).

42. The survey methodology is designed to ensure that indicators are measured cost-effectively, and with the frequency, precision and level of disaggregation that is needed to motivate for change. The data will be nationally representative, and will be repeated with a frequency of 3 years. The surveys will therefore yield repeated cross-sectional surveys not panel data. Sub-national data may be provided, largely for dialogue with sub-national audiences to give a flavor of the sub-national variation, but may not be fully representative. Any country that interested in sub-national disaggregation (e.g. very large and/or federal states) will require additional data collection and additional resources. The discussion in the ANNEX A provides the technical motivation for the key data collection decisions.

43. All data from the Program will be made publicly available and will be downloadable in a standard format from the World Bank Microdata website. This will be part of the contract with the country implementing organization. The data on the indicators themselves should be publicly available once the indicators are released. As for the other micro-level data that is collected

¹² A distinction should be made between for-profit providers in the health sector and the education sector. In the health sector, the inclusion of for profit providers may add a level of complexity, as some of the issues are beyond service delivery such as insurance, benefit package design and other issues that may create a distraction. That said, non-profit providers such as the faith-based providers should be included as they are often function alongside government providers (to the consumers) and/or offer a similar range of services.

together with the indicators, the implementing organization is allowed to restrict the publication of these data up to nine months after the release of the indicator data.

Service Delivery Indicators Trust Fund

44. As discussed, the Program will be hosted at the Bank for the first six years. The Program will capitalize on existing structures within the Bank. The Bank will administer the trust fund on behalf of the donors, serving as trustee in accordance with the legal framework outlined in the Administration Agreement,¹³ to which all contributing donors are signatories. Some of the key features of the trust fund are:

- Multidonor trust fund: Multiple donors will contribute to the pooled funds, and funders will be signatories to the Administration Agreement.
- Hybrid trust fund: This means that the trust fund is both Bank-executed (funds and activities managed by the Bank) as well as Recipient-executed (channels grants and contracts to governments or non-governmental entities).
- Programmatic trust fund: The trust fund activities are based on an agreed framework as described in the Administration Agreement. Programmatic trust funds have a two-stage execution mechanism: (i) donors agree on broad thematic framework; and (ii) grants are approved for specific activities based on criteria specified in the Administration Agreement.

45. The implementation of all trust fund activities will be subject to the Bank's operational and administrative policies and business procedures. This includes the Bank's framework on governance and anti-corruption, financial management and audits, procurement, environmental and social safeguards, as well as monitoring, evaluation, and reporting. Resources provided under the Service Delivery Indicators Trust Fund will be subject to the same control principles as other Bank resources.

46. Some of the principles that govern the governance structures and activities are:
- Inclusive governance: Framework is agreed by donors, implementing organizations and beneficiaries, and then delegated to administrative level for implementation.
 - Advisory input: Donors, key stakeholders, and technical experts provide advice at regular intervals.
 - National implementation: At the country level, the Program will be implemented through national think-tanks or other organizations in a manner consistent with the Bank's procurement guidelines.

Furthermore, the trust fund will also incorporate best practices identified in the reviews of trust funds in the Bank. Many of the lessons on the effectiveness of trust funds arise from the poor definition of the product at the time when the trust fund is created. What is unique about this trust fund is that the Program and its products are clearly defined at the outset.

¹³ An Administration Agreement is the legal agreement between the Bank and the trust fund donors. It specifies the trust fund's purposes and the scope of activities to be financed. It defines the nature of the relationship between the Bank and the donors, and mechanisms for dissolution (dissolving the trust if ever needed). It spells out arrangements governing the use of funds, progress and financial reporting, cost recovery fee, auditing, disclosure of information, among others.

V. Governance

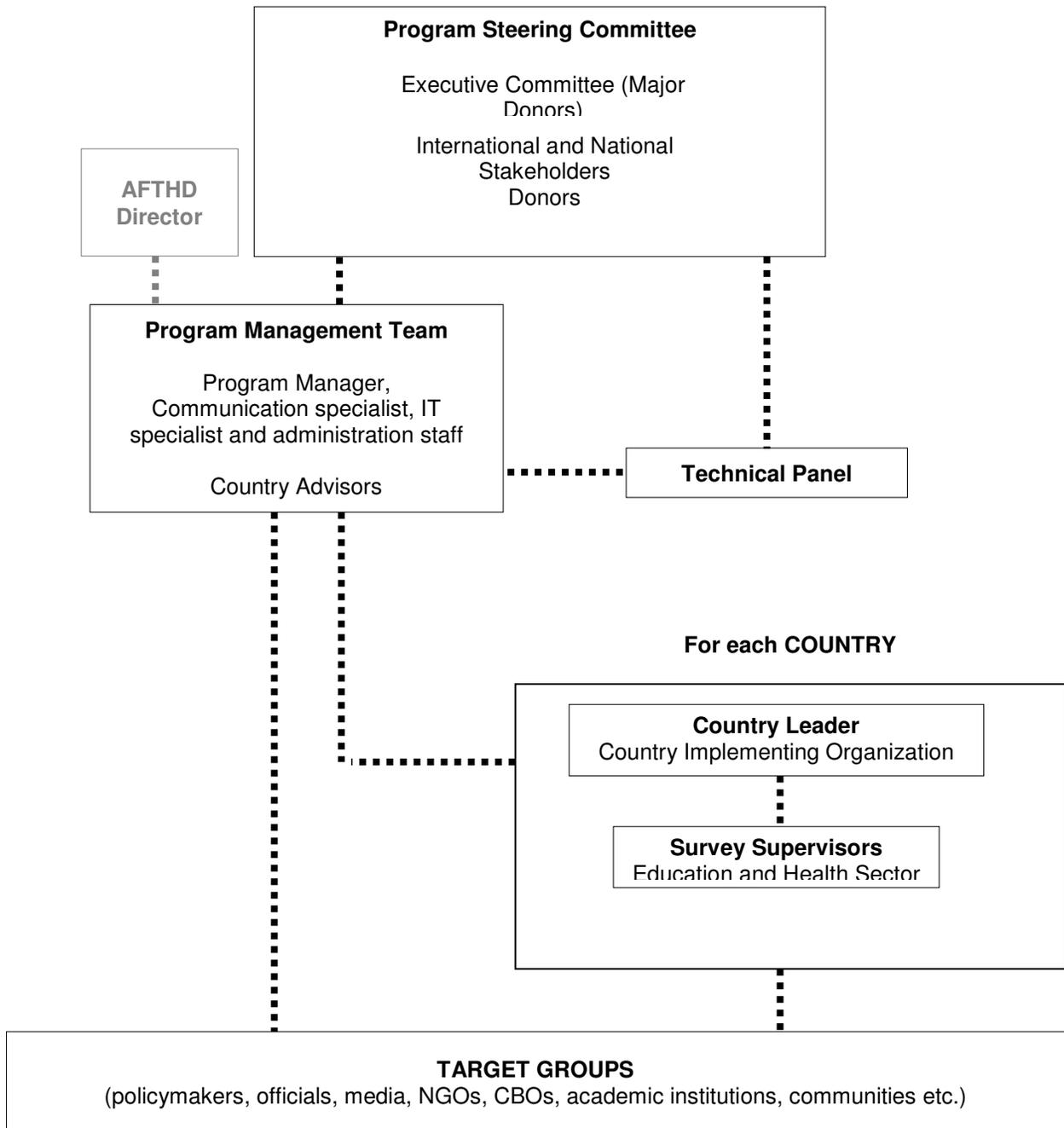
47. The Trust Fund is designed as a partnership initiative where agencies that share the vision and objectives of the Program are encouraged to participate. The Program started as a partnership initiative among the World Bank, the AERC and the Hewlett Foundation, and subsequently the AfDB joined the partnership.

Governance Structure

48. The governance structure is intended to be inclusive yet nimble, aiming to secure broad buy-in of the various stakeholders, with efficient and effective financial and technical oversight. The Trust Fund will follow the governance structure will include: Steering Committee, Executive Committee, Technical Panel, Program Management Team and Country Implementation Organization. After the first year of implementation there may be refinements in the governance structure based on the initial implementation experience.

49. **Steering Committee and Executive Committee.** The roles of the Steering Committee are to: provide guidance regarding the overall direction of the Program; promote the use of the Service Delivery Indicators in international and national policy processes; ensure high ethical standards, including high quality of data collection and integrity in the dissemination of results; and endorse annual workplans prepared by the Program Management Team. The Steering Committee will be composed of representatives of: (i) the World Bank (Africa Region); (ii) the Hewlett Foundation and other donors; (iii) AERC; (iv) the AfDB; (v) key stakeholders; and (vi) Program Manager (will serve as a Secretary to the Steering Committee). The Steering Committee members will serve a predetermined term (unless otherwise specified). The Steering Committee will have an Executive Committee that will comprise major donors (over a pre-determined threshold contribution level). The Executive Committee will: guide the overall objectives and design of the Program and shape the long-term vision of the Program; designate the Program management role to a capable institution; and approve external audits of Program accounts.

Figure 1. Program Governance Structure



50. **Program Management Team.** A Program Management team, led by a Program Manager, will: manage the SDI Program activities on a daily basis; independently implement and monitor the Program; and oversee the preparation of the SDI Report, the Annual Progress and Financial report and any special country reports as needed. A Country Advisor will be designated for each country being surveyed, and the advisor will provide support and oversight to the work of the Country Implementation Organization in each country.

51. **Technical Panel.** The panel membership is proposed by the Program Management Team and endorsed by the Steering Committee. The Technical Panel works closely with the Program structures on technical issues such as the choice of indicators, survey design and implementation, and data use, with special emphasis on quality and technical integrity of the Service Delivery Indicators.

52. **Country Implementing Organization.** The implementing organization (one per country) will be responsible for country-level implementation of the Program. The criteria for selecting a Country Implementing Organization will include: technical capacity and proven track record of collecting and analyzing survey data; demonstrated ability to work constructively with the government and other partners; and other criteria as determined by the Program Management Team in consultation with the Steering Committee. The Program Management Team provides oversight, quality assurance and technical support to the Country Implementation Organization. This support function will be executed by a Country Advisor.

53. The Program is designed as a partnership initiative where agencies who share the vision and objectives of the Program are encouraged to participate. It started as a partnership initiative among the World Bank, the AERC and the Hewlett Foundation, and subsequently the AfDB joined the partnership. Over time this partnership is expected to expand. Establishing an inclusive, yet nimble governance structure is key to securing broad buy-in of the various stakeholders, with efficient and effective financial and technical oversight.

54. The Program will be housed in the Africa region of the World Bank, in particular in the Human Development department, AFTHD, under the leadership of the Sector Director.

55. These structures roughly coincide with the Bank's levels of functions for trust funds:

- *Trustee level:* This is the level at which funds are contributed, and coincides with the function of the Executive Committee.
- *Program level:* This is the level at which funds are allocated, and coincides with the function of the Steering Committee and selected functions of the Program Management Team.
- *Grant level:* This is the level at which funds are disbursed, and coincides with the functions of the Program Management Team, the functions of the Country Implementation Organization etc. This is the level where the Bank-executed and Recipient- executed funds will be allocated.

Arrangements for Monitoring

56. The mechanisms for reporting will be:

- Financial monitoring and reporting: The Program Management Team will assess annual financial reports to secure financial governance adhering to the World Bank standards.
- Program reporting: The Program Management Team will continuously monitor the Program execution, and report to the Steering Committee at six monthly intervals, and convene an annual donor consultation forum.

VI. Budgeting and Financing

Budget

57. The resource requirements are described in the annex. Table 3 summarizes the total budget over the first six years of the Program, including a preparation phase. The budgeting assumes that five countries are surveyed each year. With a new survey implemented in each country every third year, this implies that a total of 15 countries will be covered from year 3 onwards. The total cost over the initial five year period (plus preparation phase) is estimated at USD27.7 million (see Table 3 for details about the budget estimates and assumptions).¹⁴

58. The preparation phase will commence with the following activities:

- Establish a Multi-donor Trust Fund and make necessary contractual arrangements with donors to finance the Trust Fund.
- Set up the Program Management Team (including Program Manager, Country Advisors etc.).
- Establish the Technical Panel.
- Establish the Steering Committee; convene the first meeting of the Steering Committee.
- Conduct further consultations with interested stakeholders on the Program.
- Finalize the tools for data collection based on the final list of indicators in consultation with the Technical Panel. This may include piloting, as necessary, new indicators and instruments with support from the Technical Panel.
- Agree on first round of countries to be surveyed.
- Initiate selection of the Country Implementing Organizations in the first round of countries to be surveyed.

¹⁴ The sample size and the time spent per facility / school are crucial determinants for the costs. The estimates use a sample size of 300 facilities in each sector and assume that a team of two enumerators spends 2 days per facility, on average.

Table 3. Program Budget (USD million)

Cost Component	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Preparing Implementation	1.6						1.6
Program Management Team (incl. Technical Panel)		1.8	2.1	2.4	2.4	2.4	11.2
Steering Committee		0.04	0.04	0.04	0.04	0.04	0.20
Country Implementing Organization: Annual costs		0.1	0.3	0.4	0.4	0.4	1.6
Country Implementing Organization: Survey costs		1.9	1.9	1.9	1.9	1.9	9.5
Sub-total	1.6	3.8	4.3	4.7	4.7	4.7	24.0
Contingency							3.6
Total costs							27.6

Notes: Program management team costs include: salaries (including specialist staff such as communications specialist, information technology specialist, statistician etc. communications and software needs, administrative overheads, travel, as well as costs of the Technical Panel. The assumptions used in the country-level survey estimates include: 600 facilities total (300 for education sector and 300 for health sector) to be surveyed, 2 persons per team per sector per facility, 2.5 facilities per sector per week i.e. 5 facilities per week. Data collection over period of 12 weeks by 10 teams per sector i.e. 20 teams and 40 persons in total. This includes 10 supervisors and 30 enumerators. Five countries will be surveyed per year, and 15 countries included in total over the Program period. A new survey in each country every third year.

Financing

59. A Multi-donor Trust Fund will be established. The donors that have expressed strong interest are: the Hewlett Foundation, NORAD, DFID and Finland. Proposals are being submitted to these potential funders. The Bank will be opportunistic in seeking additional funding sources, although participation as a donor will be subject to fully respecting the methodology and protocols of the Service Delivery Indicators Program. Furthermore, the conditions specified in the Administrative Agreement governing the Multidonor Trust Fund will be applied to all incoming donors, as per the Bank's Trust Fund guidelines.

60. Resources may be mobilized from IDA funded Programs within AFTHD. Consistent with Bank operational requirements, this will be subject to the government's full agreement, the Bank's procurement processes and oversight by relevant staff, who are ultimately accountable to the Director, AFTHD.

ANNEX A. TECHNICAL ANNEX

Data Collection and Survey Methodology

1. All data will be collected at the health facility or school level. In the education sector, focus will be on primary schools, while in the health sector to focus will be on primary health services up to the level of first referral hospitals (outpatient services and maternal and newborn services). Government, non-profit and private for-profit facilities will be included as appropriate. The data will be gathered from direct observation of provider behavior, from various test of provider knowledge and skills, and from observation of the availability of key inputs required to enable provision of quality services (infrastructure, equipment, supplies, etc.).
2. The pilot developed detailed manuals and data collection tools for the indicators. New indicators, and those that have been revised, need to be piloted before implementation at scale. Data will be collected by trained enumerators, one team for each sector (education and health) supervised by a survey supervisor, again under the supervision of the country leader within the implementing institution in each country.
3. The survey methodology is designed to ensure that indicators are measured cost-effectively, and with the frequency, precision and level of disaggregation that is needed to motivate for change. At least three issues are important in this context:
 - Level of aggregation (figures reported at national or sub-national levels)
 - Repeated surveys with the same participants (panel data) or new participants in new survey rounds (repeated cross-section data)
 - Frequency of surveys

Level of Aggregation and Sample Size

4. It is suggested that the surveys are designed with the aim of producing nationally representative indicators with sufficient precision to identify changes in the indicators of around 5-7 percentage points over time. At the same time, separate results will be reported for urban/rural areas and government/non-government providers where appropriate. The sub-national aggregates will have lower levels of precision.
5. In order to minimize costs, the normal approach to sampling will be a multistage, cluster sampling approach, for instance by first selecting districts (or another suitable geographic unit) and then selecting health facilities and schools within the selected districts. Several of the indicators will also involve a third step by selecting teachers / health workers within the facilities. Cluster sampling generally increases the variability of the sample estimates above that of simple random sampling. Hence, the number of surveyed units will increase. These costs must be weighed against the added costs of travel and administration with a simple random sample.
6. Based on the results from the pilot, we have tentatively estimated that a nationally representative sample with cluster sampling would require at least 250-300 health facilities and

the same number of schools in each country.¹⁵ (The pilot used samples of 150-175 units in each sector). Adjustments to the sample size should be routinely considered as new data are collected and analyzed.

7. The level of disaggregation has large implications for sample size. Since sample size calculations apply to the smallest sub-group represented, a basic rule of thumb is each time data is disaggregated into two sub-groups, the required sample size is doubled. Thus, in order to make the rural/urban distinction with the same level of precision as the national aggregate, we need twice the sample size as what would be necessary for national level estimates.¹⁶ This “rule” assumes that the total population is “large” (i.e., in the thousands), even at the disaggregate level. This assumption is not necessarily valid when our populations consist of schools, health facilities and service providers. But even when we take into account the implications of smaller populations (finite population corrections), the implications for sample size of moving to lower levels of aggregation are large.¹⁷

8. Our consultations with national stakeholders in Zambia, Ghana and Kenya suggest that the Service Delivery Indicators represent a forceful tool when reported as a national aggregate, but that it may also be desirable to report the Indicators separately for urban/rural settings and perhaps also for smaller geographical areas (region/district).

9. The proposed level of reporting balances these benefits and costs. The most critical issue is to monitor progress at national level over time. The fact that sub-national indicators will be measured with lower level of precision is regarded as a minor problem as the purpose of distinguishing between rural/urban areas and government/non-government providers usually will be to identify any large differences between them.

10. When it comes to non-government providers, our recommendation based on consultation with various stakeholders is that the indicators should reflect both government and private non-profit providers, but not private-for profit providers. However, in some countries, private providers play a prominent role in especially in primary education. To have the possibility to compare the quality among these providers with the quality in the not-for-profit sector could be an important piece for information to motivate for change. It is therefore suggested that, where appropriate, the school survey adds private-for-profit providers to the sample. The purpose is to obtain a benchmark and to identify any large differences between not-for-profit and for-profit providers, and not necessarily to follow the progress of private-for-profit providers over time. We therefore suggest that the sample of private-for-profit providers should be relatively small (e.g. 1/3 of the not-for-profit sample).

¹⁵ Assuming 80% power, 10% level of confidence, and a survey design effect of 2.

¹⁶ If variation is lower within the sub-groups than in the total sample, the total sample size is somewhat reduced. However, the pilot results indicated that the levels of variation within the urban/rural strata are not much smaller than in the national aggregate.

¹⁷ Example: Assume a country has 100 districts with 50 health facilities or schools in each. Further assume that we want a level of precision that requires a sample of 200 units from a population of 5000 when we want results to be reported at the national level. If we instead decided report at the district level, the finite population correction may reduce the required sample to 40 of the total population of 50 in each district, but this would still result in a total population of 4000 units in total. Hence, the sample size increases by a factor of 20.

Repeated Cross-section versus Panel Survey

11. Two approaches are available for identifying change in the indicators over time. One is to collect data from the same units (and possibly the same service providers) in consecutive survey rounds (panel data). The other is to draw a new random sample every time the survey is conducted (repeated cross-section). We recommend using the latter approach here.

12. The benefit of panel data is that there will be less “noise” in the data, implying that a smaller sample size is needed in order to identify change over time. The degree to which the panel would reduce the sample size requirements depends on the individual variable and the degree of correlation between years. For example, a correlation of 70 percent would reduce the required sample size by 30 percent in each round.

13. The panel data technique has several drawbacks, though. One potential problem is that once it is known which units will be surveyed, special efforts can be made to improve the performance of these units. Another challenge is that there may be learning effects from participating in the survey. Hence, the panel may evolve into a biased sample over time. Moreover, attrition of service providers may make it difficult to trace the same individuals, effectively reducing the benefits of the panel structure.

14. In this case, though there would be benefits to panel data in terms of sample size and analytical opportunities, the probable bias resulting from knowing that one is monitored over time would probably outweigh the benefits.

Frequency of Surveys

15. The frequency of the surveys should be determined based on what would be most effective in motivating for change. On the one hand, frequent repetition of the surveys is important to maintain a continuous focus on the indicators and on quality in service delivery. On the other hand, several of the indicators are unlikely to change much from year to year.

16. We believe that a three-year sequence is sufficient for measuring change, because many of the indicators are difficult to move much over shorter periods of time. From an advocacy perspective, however, three-year intervals do not seem sufficient. Experience from the Doing Business Index suggests that one survey cannot be expected to maintain public interest over more than a nine month period, even if large efforts are made to this end. It therefore seems important to have new results out every year.

17. To reconcile these considerations, the following approach is suggested: New surveys are implemented every third year in each country, but the surveys are sequenced across countries such that one third of the countries are surveyed in year one, another third in year two, and so on. Hence, new data will be coming in every year. A new international report will then be produced every year, partly based on new data and partly based on data from the two preceding years (the last survey in all countries). In this way, there will be new stories to tell every year about the relative performance across countries. This approach also makes it possible to use the

international reports to build expectations to the countries (e.g. “next year, new data from your country will be reported, will you be able to report progress?”)

Data Analysis and Availability

18. Several possibilities exist for facilitating further analysis and use of the data by research institutions and think-thanks. One possibility is to adapt the ADePT software tool developed by the research department of the Bank to facilitate analysis and presentation of the survey data¹⁸. This tool will be available for downloading and will be especially useful for the country implementing organizations in their production of reports and other dissemination outputs. Another possibility would be to create online tools like the ones used by the DHS surveys (STATcompiler and STATmapper). The current budget includes adaptation of the ADePT tool, which is more flexible than the online tools, but the appropriateness of the alternative solutions need to be further discussed before a final decision is made.

19. All data from the Program will be made publicly available and will be downloadable in a standard format from the SDI web page. This needs to be part of the contract with the country implementing organizations. The data on the indicators themselves should be publicly available once the indicators are released. As for the other data that is collected together with the indicators, we suggest that the implementing organization is allowed to restrict the publication of these data up to nine months after the release of the indicator data.

Education Indicators

20. The education service indicators provide a snapshot of the learning environment and key set of resources, including human resources, which need to be in place for students to learn. A strong focus is placed on the knowledge, skills and effort of teachers. This is motivated by several facts. Expenditure on teachers represents by far the largest share of education spending in developing countries. Moreover, many recent studies demonstrate how changes in teacher behavior can improve learning achievement.¹⁹ Despite this, little is known about the capabilities of teachers and the quality of teaching in developing countries. How much time do they spend teaching? How well do they know the subjects they teach? How effectively do they conduct their lessons?

21. A minimum requirement for learning is that the teachers are not *absent from the school* and spend time in the classroom rather than somewhere else (*absence from the classroom*). However, while having teachers in the classroom is a necessary condition, it is not sufficient. Teachers need to have at least a *minimum level of knowledge* of the subjects they are teaching and exert a certain level of effort to enable learning. An important element of the indicators is the attempt to directly measure teacher effort in the classroom, which has not been done consistently to date. One indicator therefore measures whether teachers stay in the classroom during the

¹⁸<http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTPROGRAMS/EXTADEPT/0,,contentMDK:22595675~menuPK:7108374~pagePK:64168176~piPK:64168140~theSitePK:7108360,00.html>

¹⁹ See, for example, E. Duflo, P. Dupas and M. Kremer (2009) “Additional Resources versus Organizational Changes in Education: Experimental Evidence from Kenya”, MIT, mimeo, May (and the references therein).

entire lesson and spend their time on teaching activities (*time spent teaching in the classroom*). Another indicator assesses the extent to which the teachers' teaching methods are such that students absorb the lesson (*quality of instructions*).

22. Two indicators will assess if sufficient materials are in place to support the teaching activities (*minimum teaching equipment available* and *textbooks per student*), and another indicator measures the *student-teacher ratio*. Finally, the level of school infrastructure will be assessed, with a particular focus on the availability of sanitation facilities and sufficient light in the classroom.

23. Below we present in more detail the nine indicators highlighted above and how they will be measured. All indicators are based either on tests or derived from direct observation by trained enumerators.

Indicator E1: Absence from schools

24. The indicator is defined as the share of teachers in schools as observed during one unannounced visit. It is constructed as follows: During the first (announced) visit, 10 teachers are randomly selected from the list of all teachers, and the whereabouts of these 10 teachers are then verified in the second, unannounced, visit. The number of pre-identified teachers still on the roster who are in school (but not necessarily in a class room) during the unannounced visit is then divided by the number pre-identified teachers still on the roster, to give the rate of absenteeism.

Indicator E2: Absence from classroom

25. The indicator is defined as the share of teachers who are present in the classroom during scheduled teaching hours as observed during an unannounced visit. It is constructed in the same way as indicator E1, and presented as absence rate from classroom, with the exception that the nominator now is the number of pre-identified teachers still on the roster who are both in school and in the classroom, based on the unannounced visit.

Indicator E3: Share of teachers with minimum knowledge

26. The indicator assesses whether primary school teachers have sufficient professional knowledge to:

- (a) master aspects of primary mathematics, such as manipulation of numbers, place values, time, measurement, and fractions.
- (b) master aspects of primary language teaching such as reading comprehension and grammar.
- (c) read a variety of simple information-giving texts and extract and summarize information.
- (d) mark children's work in language and mathematics.
- (e) add up marks on class tests, turn raw scores into averages and percentages, and read and comment on bar charts containing information on children's learning achievements so as to monitor the learning progression of individual or groups of students.

27. The test measures subject knowledge and pedagogical content knowledge. We propose to test all (a maximum of 10) grade 3-4 teachers in the subjects that they teach. Minimum

knowledge is defined as scoring 80% or more on the combined test.²⁰ We focus on teachers teaching younger cohorts because cognitive ability is most malleable at younger ages (see Cunha and Heckman, 2007), therefore we would expect that teacher effort has the highest marginal effect at early age.

Indicator E4: Time spent teaching in the classroom

28. The indicator measures the amount of time a teacher spends on teaching (vs. non-teaching activities) when giving a lesson, measured as the share of a lesson devoted to teaching.

29. We distinguish between teaching and non-teaching activities based on classroom observation done inside the classroom. We define teaching very broadly, including actively interacting with students, correcting or grading student's work, asking questions, testing, using the blackboard or having students working on a specific task, drilling or memorization, and maintaining discipline in class. We define non-teaching as work that is not related to teaching, including working on private matters, doing nothing and thus leaving students not paying attention, or leaving the classroom altogether.

30. The presence of the enumerator inside the classroom will presumably make the teacher less likely to leave the classroom than they would in a normal teaching situation. The share of the time spent outside the classroom will therefore be measured through an alternative approach; based on classroom observation done outside of the classroom we will record during a 15 minute interval whether the teacher is present in class or not (record minutes of absence). This gives a snapshot of how many teachers are outside the class room when they should be teaching a lesson.

31. The indicator will be calculated by first measuring the time used on teaching activities based on classroom observation done inside the classroom and then adjusted by the time spent outside the classroom based on the 'outside' classroom observation.

32. To give an example: First, suppose the teacher is present in the class room for 7.5 out of the 15 minutes. Second, suppose the teacher is observed to spend 22.5 minutes of a 45 minute lesson on teaching activities during the inside class room observation. Then we would conclude that the teacher spends $22.5/45 \times 0.5 = 25\%$ of every lesson teaching.

33. We propose to observe two mathematics lessons and two language lesson in grade 3-4. The first observation (of two classes) could take place on the day of the main visit and the second observation during the unannounced visit.

Indicator E5: Quality of instructions

34. This indicator attempts a direct measure of the quality of instruction pupils receive, based on classroom observations. Specifically, the enumerators will identify the subject, the topic, and the key instructions given to students. During the class, 10 random students will be observed

²⁰ Data from the pilot suggest that scoring the level of knowledge as a 0/1 variable rather than as a continuous variable implies a much larger standard deviation and therefore require a much larger sample size. This issue needs further consideration, both here and for other, similar variables, as more data are collected.

every 5 minute to see if they are engaged in learning activities. After the class, the same students will be asked one-by-one whether:

- (a) they know the subject of the lesson
- (b) they know the topic of the lesson
- (c) they know what they were asked to do in a learning activity that took place during the lesson.

35. The indicator is calculated by first measuring the share of the time that students are engaged in learning activities based on the classroom observation of 10 students. This share is then added to the shares of students that have correctly understood (a), (b), and (c), and the aggregate is then divided by 4 (thus giving equal weight to each of the four components). The great advantage of this approach is that we do not have to assess and evaluate how a teacher delivers the lesson, since this could be done in a variety of different ways that may be difficult to compare and aggregate. Instead, we simply assess whether the lesson has achieved its goal.

Indicator E6: Availability of teaching resources

36. This indicator measures whether teaching equipment is available in the classroom for the teacher to teach and students to learn. The list of equipment needed at a minimum is: blackboard, chalk, pencils, paper. (We will also record whether desks are present, and this could potentially also be included in the indicator, although desks are not necessarily needed for quality teaching.)

37. The indicator will be measured as follows: we assign a score of 1 if a blackboard and chalk are present. Then we calculate the share of students who have pencil and paper available. We add these scores and divide by the maximum possible score to arrive at the % of minimum teaching equipment available. Alternatively, the indicator is binary and measured as minimum teaching equipment available or not. In this case, we assign a score of 1 if there is a blackboard, chalk, and every student has pen and paper.

Indicator E7: Student-teacher ratio

38. Student-teacher ratio is measured as the average number of students per teacher. It is based on the classroom observation schedule, where we simply count the number of students per teacher teaching.

Indicator E8: Textbooks per student

39. The indicator is measured as the number of mathematics and language books that students use in a grade 4 classroom divided by the number of students present in the classroom. The data will be collected as part of the classroom observation schedule.

Indicator E9: School infrastructure

40. School infrastructure is measured as the share of schools with sanitation facilities that are used by girls and sufficient light in the classroom so that students can read and study.

41. The indicator is binary. We assign a score of 1 if there are sanitation facilities that are used by girls (assessed through direct observations) and if the classroom has at least 800 lux based on test with a mobile light meter in the classroom. If one of these conditions does not hold, we give a score of 0.

Health Service Indicators

42. The health service indicators are similarly selected to portray a picture of the health services, as experienced by the users, at health facilities up to the level of first referral hospitals²¹. The main focus is on the knowledge, skills, and efforts of health workers. In addition, the Indicators assess whether the working environment is conducive to the delivery of quality services. The availability of qualified and motivated health workers is fundamental to the delivery of quality health services. However, our knowledge of what health workers know and what they do in their jobs is limited. How much time do they spend providing services to patients? Do they have the knowledge and practical skills required to detect severe illness and prevent unnecessary deaths? Do they make the efforts needed to put knowledge into practice?

43. For the patients to achieve adequate health services, a minimum requirement is that health workers are not *absent from the health facilities*, that they have the knowledge and skills required to diagnose common illnesses and prevent unnecessary deaths (*skills to reach correct diagnosis* and *skills to handle life-threatening complications for newborns and mothers*), and that they make an effort to apply this knowledge when they are seeing patients (*adherence to clinical guidelines*).

44. Health workers also need basic supplies and equipment to support their practice, such as thermometer, stethoscope, blood pressure machine, etc. (*availability of medical equipment*), and they need *drugs in stock*. In addition, sufficient time should be available for each patient consultation (*workload per clinician*). Finally, the indicators include information about the presence of key health facility infrastructure, such as sanitation facilities, amenities that protect the patients from rain and sun, and a functioning refrigerator at facilities that store vaccines (*health facility infrastructure*).

45. All indicators are based on observation and/or tests. The proposed indicators are further described below.

Indicator H1: Absence from the health facility

46. The indicator is measured as the share of health workers absent at health facilities, as observed during one unannounced visit. It is constructed as follows: During the first (announced) visit, 10 health workers are randomly selected from the list of all health workers, and the whereabouts of these 10 health workers are then verified in the second, unannounced, visit. The number of pre-identified health workers still on the roster who are present during the unannounced visit is then divided by the number pre-identified health workers still on the roster, to give the rate of absenteeism.

Indicator H2: Skills to reach correct diagnosis

This indicator is measured as the share of a set of hypothetical patient case scenarios that the health worker is able to diagnose correctly.

²¹ A first referral hospital is the lowest hospital level point of contact for patients, either as a first contact point or when referred from a lower level facility. In many low-income countries, a large share of patients uses these hospitals for primary health care.

47. Health workers who consult patients (i.e., prescribers) in the outpatient department will be presented with five different hypothetical patient case scenarios (clinical vignettes²²). The health worker is provided with a few core symptoms and is then asked to proceed to diagnose the hypothetical patient by taking the patient history and by examining the patient. The vignettes will be constructed such that the symptoms of the patient, taken together, provide a clear picture of what the patient suffers from. Hence, if the health worker asks the questions and performs the examinations he/she is supposed to, and if he/she has sufficient knowledge to interpret this information, the health worker should be able to easily reach a correct diagnosis.

48. The set of diagnoses presented to the health workers should represent a broad spectrum of the conditions that the health worker may encounter. It should include diagnoses that represent a high burden of disease. At the same time, it should represent illnesses that are less common but nevertheless severe in order to test whether the health worker is able to distinguish between the common and the less common conditions. It should also include a chronic illness. Finally, in order to ensure comparability across countries, it should only include illnesses that are prevalent in the whole region. We propose to include four conditions; pneumonia and diarrhea in children (the biggest causes of child mortality),²³ meningitis (testing the ability to detect a less common, but very severe, condition), and diabetes (representing the rapidly increasing burden of non-communicable disease).²⁴

49. The indicator is calculated by assigning a score of 1 for each correct diagnosis and then dividing by the maximum number of correct diagnoses.

Indicator H3: Ability to save the life of newborns and their mothers

50. This indicator measures health workers' skills in handling life-threatening complications of newborns and their mothers. It will be measured only at facilities that provide delivery services, and the test will be administered to those who are on duty to provide delivery services on the day of observation. The test will be performed by applying two newly developed simulation tools, presently used for teaching purposes for maternal and newborn health services in Africa and elsewhere. The simulation equipment consists of one tool (MamaNatalie®) that simulates a mother who gives birth. The tool will be used to simulate post-partum hemorrhage, the most important cause of maternal death. Another tool, NeoNatalie®, is a doll that simulates a newborn baby. NeoNatalie will be used simulate breathing problems (asphyxia), the most important cause of neo-natal death. (The tools also allow for simulation of other severe maternal and newborn conditions.)

51. For each of the complications, there is a set of procedures that the service providers are supposed to adhere to. The indicator is calculated by assigning a score of 1 for each correct procedure taken to save the baby/mother and dividing by the total number of procedures that should be adhered to.

²² Clinical vignettes represent a validated methodology for measuring quality of health services (see J.W. Peabody, M.M. Taguiwalo, D.A. Robalino, and J.Frenk (2006) "Improving Quality of Care in Developing Countries", in D.T. Jamison et al. *Disease Control Priorities in Developing Countries* (2nd edition).

²³ R.E. Black et al. (2010) Global, regional and national causes of child mortality in 2008: a systematic analysis. *Lancet*, 375(9730), 1969-87.

²⁴ Africa's Neglected Epidemic: Multidisciplinary Research, Intervention and Policy for Chronic Disease. *British Academy*, May 2010.

Indicator H4: Adherence to clinical guidelines

52. This indicator measures the share of procedures prescribed by clinical guidelines that health workers adhere to while consulting patients. The indicator is measured through direct observation of patient consultations in outpatient departments.

53. The procedures surveyed will be a generic list of items essential to quality care for patients with common infectious and chronic diseases. The list will depend on the presenting symptom (fever, cough, diarrhea, weakness/dizziness) and will include important history taking questions and physical examinations (e.g. taking temperature and measuring blood pressure).²⁵

54. For children, the list of essential procedures will include core items from the protocol for Integrated Management of Childhood Illness (IMCI), a tool developed by the WHO and UNICEF for accurate identification of childhood illnesses in outpatient settings in low-income countries, and for appropriate combined treatment of major childhood illnesses.

Indicator H5: Availability of medical equipment and supplies

55. This indicator measures the share of basic equipment and supplies that are available at the health facility. The list of essential equipment includes items such as thermometer, stethoscope, blood pressure machine, etc. For facilities that provide delivery services, it will also include items necessary for safe delivery and newborn care.²⁶

Indicator H6: Drugs in stock

56. This indicator measures the share of essential drugs in stock on the day of observation. The list of drugs will include drugs needed to treat malaria, pneumonia and other serious infections, diarrheal disease and dehydration, food supplements, vaccines, birth control measures, etc. At facilities providing delivery services, the drug list will also include important drugs needed to provide lifesaving maternal and newborn care. (The list of drugs is derived from the same WHO sources used to construct H5, see footnote).

Indicator H7: Workload per clinician

57. Actual time spent per patient will be recorded during direct observation of outpatient consultations over three hours. Through observations, the number of treated patients per clinician (per day) will also be recorded. By combining these two sources of data, one can calculate the aggregate workload per clinician (in minutes per day).

Indicator H8: Health facility infrastructure

58. Infrastructure quality will be measured for each health facility as the share of the following items that are available:

- a) Functioning toilets for both male and female patients
- b) Functioning refrigerator (at facilities with storage of vaccines)
- c) Waiting area that protects patients from sun and rain

²⁵ Oxford Handbook of Tropical Medicine. Third edition. Oxford University Press. UK. 2008.

²⁶ The list of equipment and supplies is derived from the *WHO Packages of interventions for family planning, safe abortion care, maternal newborn and child health* (2010) and from the WHO template for *Service Availability Mappings* (SAM).

- d) Seats and space to accommodate all patients in the waiting areas (based on random checks throughout the day of observation).

59. Alternatively, this indicator may be scored as 1 when all items are available and zero otherwise. All items will be assessed based on direct observations.

Additional Data on Explanatory Factors

60. The data collection effort outlined above will also provide additional information (apart from the indicators), that can be used to assist the interpretation of the indicators, and to identify potential actions for improvement.

61. First, disaggregation of the data underlying each of the indicators will provide a wealth of useful information about potential areas for improvement in terms of knowledge, skills and key inputs, and about which aspects of performance are more or less critical to further strengthen. For example, by disaggregating the information used to measure “Share of teachers with minimum knowledge”, the data collected would not only tell us whether teachers’ have sufficient knowledge to teach, but more detailed analysis can reveal which areas specifically need improvement. For the health indicators “Skills to reach correct diagnosis”, “Skills to handle life-threatening complications for newborns and mothers”, and “Adherence to clinical guidelines” a similar disaggregation can easily be done to inform policymakers about key areas where improvements are needed. Furthermore, by comparing what health workers do when their skills are tested in the clinical vignettes with what they do with real patients, we may gain important insights about the importance of knowledge and skills compared to effort in explaining actual performance. A short cognitive knowledge test of health workers, similar to the teacher test, will also be included in the data collection to further assist the interpretation of these data.

62. Second, the underlying data will shed light on where the potential for improvement is largest. As part of the data collection effort, a set of important correlates will be collected, such as access to electricity and water at the school and facility, characteristics of the staff, and delays in salaries. This will make it possible to identify, for example, what type of teachers or health workers (for instance with respect to formal education) are the most likely to be absent, or if and to what extent salary delays are correlated with staff morale and therefore with the quality of service.

63. Third, with little extra costs, we may collect data on additional factors that are important for the quality of service delivery, but are not included among the Service Delivery Indicators, either because they are difficult to measure on an ordinal scale, difficult to aggregate into a quantitative score, or can be seen as underlying causes of behavior rather than as primary inputs in service production. Some of these are related to efforts at the health facility level, others are related to efforts higher up in the supply chain. One aspect is whether performance standards, such as opening hours and patient rights, are clearly defined and communicated. Furthermore, an extensive literature in sociology and managerial science has highlighted the role of adequate managerial and leadership skills to organize and promote quality routines and behavior. Similarly, opportunities to provide feedback on the level of quality and whether action is being

taken in cases of substandard quality are important in the assessment of service quality. While these aspects are difficult to measure, it is possible to at least collect proxy measures for them.

64. Furthermore, from the classroom observation schedule, we can derive valuable additional information about teaching practices and teacher's ability to translate their professional knowledge (knowledge of content and pedagogy) into teaching in class. The classroom observation schedule approach follows Johnson (2006) in that the schedule divides a lesson in three different phases: introduction, main body of the lesson, and conclusion, and checks whether teachers use techniques appropriate for each part of the lesson. This information does not form part of the indicator, because it is difficult to aggregate teaching methods into a quantitative index. Nevertheless, it serves two useful purposes: (i) by linking it to the "Quality of instructions" indicator one can assess which teaching methods proved most successful in engaging students and transferring knowledge; and (ii) it provides useful information to policy makers when designing in-service training programs.

65. On efforts higher up in the supply chain, it would be straightforward to add a module to measure the efficiency of the supply chain, by estimating resource leakage in support programs for either primary schools or primary health facilities. However, as the type of support programs differ across countries, this exercise should rather be done on a country-by-country basis.

66. Fourth, the pilot demonstrated that it is possible, with little extra effort, to measure outcomes in education. We do not suggest that outcomes should be part of the Service Delivery Indicators, because the link between inputs, performance and outcomes is complex and to a considerable degree beyond the control of the service providers. However, to collect outcome data (in selected countries) might still be of great value to the Program, because it could be used to illustrate the close link between the Service Delivery Indicators and outcomes (as was done in the pilot) and thus strengthen the cause of the Program. Data on learning outcomes over time would also be very useful for monitoring purposes.

67. Finally, the survey could be expanded with modules on various country-specific aspects that policy makers and other stakeholders have identified as important.

68. Several of the modules needed to collect the additional information described above were included and tested in the pilot. Other modules need to be developed in consultation with the technical experts.

ANNEX B. ORGANIZATIONAL RESOURCES REQUIRED (PRELIMINARY)²⁷

Program Management Team

1. The resources described are the estimated requirements for implementing the Program activities for the first wave of five countries.

Program Manager

2. A full-time Program Manager will have extensive professional training and knowledge (preferably at doctoral level) within the fields of health and/or education economics with at least 10 years of experience in a developing country context. Additional qualifications include excellent negotiation and advocacy skills in a multicultural environment.

Country Advisors

3. One country advisor will cover three countries. These country advisors will have extensive professional training and knowledge including survey design and implementation (at least at Masters Level) with at least 3–5 years of management, training or consulting experience in a developing country context. Additional qualifications include strong interpersonal communication and capacity building skills.
4. Each advisor will initiate and closely monitor data collection in three countries. Although the implementation will be contracted to and coordinated by country implementation organizations, there is a need for on-the-ground, close follow up in terms of implementing the contract, providing support, and building capacity on a continuous basis. To follow up the implementation of the data collection, analysis and extensive dissemination will require physical presence or frequent travel.
5. When the Program is fully implemented in 15 countries, each Country Advisor will manage one country survey and do analytical and dissemination support in two additional countries each year. During the two first years of the Program, while the number of countries has not yet reached 15, survey implementation will occupy a larger share of the Country Advisors' time given learning by doing. This seems reasonable as the first implementation will be more labor intensive both in terms of sensitization of stakeholders, capacity building of implementing agencies, designing and writing the first country reports, identification of suitable dissemination channels and formats, etc.
6. The Country Advisors will report to the Program Manager.

Other staff resources

7. Other staff resources needed in the Program Management Team are:
 - Communication expert (100%)
 - Administrator (100%)
 - Statistician (50%)

²⁷ These roles are preliminary and will be finalized within the first year of Program operation.

- IT/Web coordinator (20%)

Country implementing Organization

Country leader

8. The country leader will be responsible for the implementation of the Program as well as coordinating Program preparations, data collection, data analysis, and dissemination activities at the country level. This role will be required for 2-3 months per year depending on Program cycle. The country leader will also be responsible for ensuring that all logistics, permits etc. have been arranged. Country leaders should have a minimum of 3-5 years of survey implementation experience with particular emphasis on Program planning and data analysis skills.

Survey Supervisors

9. The survey supervisors will leading the implementation of the survey in each sector, with responsibility for quality assurance of the data, training and coordination of enumerators at headquarters and in the field and outreach and dissemination activities with assistance from the country leader. This position will be a full time position during the survey year to prepare and implement the survey, analyze the data and prepare country reports, and part time during the two dissemination years to continue analyzing and interpreting the underlying data and disseminating the results. The survey supervisors for health will require a university degree in either medicine or nursing and the supervisors for education will require a relevant degree in education or economics. All survey supervisors will have basic research training (statistics, data collection, data analysis and strong writing skills).

Enumerators and Data Entrants

10. Enumerators will work together in teams of two to collect the data, spending two days at each health facility or school. The number of teams will depend on the sample size and the total time allocated to data collection. The current budget assumes that 40 enumerators (including 10 field supervisors) are needed over a 12-week period to collect the data from two sectors (300 facilities in each sector). Enumerators will be supervised both by the education and health survey supervisors and by the country leader.

ANNEX C. INFORMATION PERTAINING TO DATA ACCESS

World Bank Open Data Policy

According to the Bank's Open Data Policy:

- Users are free to copy, distribute, adapt, display or include the data in other products for commercial and noncommercial purposes at no cost subject to certain limitations summarized below.
- User must include attribution for the data you use in the manner indicated in the metadata included with the data.
- Users must not claim or imply that the Bank endorses your use of the data by or use The World Bank's logo(s) or trademark(s) in conjunction with such use.
- Other parties may have ownership interests in some of the materials contained on The World Bank Web site. For example, the Bank maintain a list of some specific data within the Datasets that you may not redistribute or reuse without first contacting the original content provider, as well as information regarding how to contact the original content provider. This is accommodated in the terms of use.
- The World Bank makes no warranties with respect to the data and users agree The World Bank shall not be liable to you in connection with your use of the data.

Note: This is only a summary of the Terms of Use for Datasets Listed in The World Bank Data Catalogue.

World Bank Microdata Website (<http://microdata.worldbank.org/>)

Data disseminated by the World Bank fall under the Open Data policy. A number of concerns prevent the World Bank from providing similar unconditional and unrestricted access to microdata.

Microdata are unit-level data obtained from sample surveys, censuses, and administrative systems. They provide information about characteristics of individual people or entities such as households, business enterprises, facilities, farms or even geographical areas such as villages or towns. One main concern is the need to protect, as far as possible, the privacy of the individual respondents that have provided the data. Much of the microdata have been collected under a mandate provided by different kinds of national statistical legislation, which contain requirements for data about individuals to be treated as confidential by anyone collecting and using such data. Further, the microdata may be protected by copyright or other intellectual property protections under different intellectual property legislation. The Microdata Library adopts rules and procedures that are consistent with the type of legislation concerning this information around the world.

The Microdata Library operates as a portal. It disseminates micro-datasets from two sources: those that belong to the World Bank and those where the data have been generated and are owned by another agency. A number of micro-datasets have been provided to the Microdata Library by third parties including member states, international or regional agencies, and World Bank contractors. In many cases the arrangements that have allowed the datasets to be made

available to the World Bank include limitations on how the data can be disseminated to users. The Microdata Library is obliged to put these arrangements into effect.

For these reasons, the Microdata Library disseminates data under some restrictions, the terms of which vary by source, as well as according to the individual characteristics of each microdata set. While the Microdata Library remains governed by the Terms of Use for the World Bank's datasets, it is subject to additional terms, described below on the Microdata website (<http://microdata.worldbank.org/>). If these additional Microdata terms are inconsistent with the general Terms of Use, then these specific Microdata terms shall govern.

ANNEX D. CONTRIBUTORS AND PROGRAM PREPARATION TEAM

The Program Document is largely based on the document prepared by the Chr. Michelsen Institute (CMI) with generous funding from the Hewlett Foundation. The contributing authors include Ottar Mæstad (CMI), Øystein Evjen Olsen (independent consultant), Jakob Svensson and Tessa Bold (Institute for International Economic Studies (IIES), Stockholm University), Bernard Gauthier (HEC Montréal), and Mwangi Kimenyi (Brookings Institute and AERC). Ritva Reinikka (World Bank) provided strategic guidance during the process. Important inputs were received through consultations with education and health stakeholder in Ghana, Zambia and Kenya, coordinated with the assistance of Charles Michelo (University of Zambia), Charles Dollie (Africa Center for Economic Transformation (ACET), Ghana), James Nduko (Twaweza, Kenya) and Sara Ruto (Uwezo, Kenya). Valuable advice and inputs have also been received from a number of individuals; Kristen Himelein, Kaliope Azzi-Huck, Kavita Watsa, Jerry Lebo, Shireen Mahdi (World Bank), David Johnson (Oxford University), William Howlett (KCMC Tanzania), Sven Gudmund Hinderaker, Erling Svensen, Nils Gunnar Songstad, and Odd Mørkve (CIH, University of Bergen), Joseph Dwyer (MSH), Karen Caldwell (MSH, Kenya), Onyango Ouma (IAGAS), Daraus Bukunya (MSH) Olu Ajakaiye (AERC) and Hege Langli Ersdal, (SUS, University of Oslo). Inputs have also been received from NORAD and DFID. An earlier draft of this document was reviewed in a consultative meeting in Bergen in July 2011, with participants from Hewlett Foundation, World Bank, AERC, AfDB, CMI and NORAD.