



# **PFM Performance Measurement Framework**

## **Monitoring Report 2009**

*for the period April 2007 – March 2009  
with assessment status statistics as of October 2009*

**PEFA Secretariat  
Final**

**February 2, 2010**



# Table of Contents

<b>List of Abbreviations</b> .....	<b>4</b>
<b>Executive Summary</b> .....	<b>6</b>
<b>Chapter 1 Introduction</b> .....	<b>9</b>
<b>Chapter 2 Overview of Application of the Framework</b> .....	<b>10</b>
2.1 The rate of roll-out of the Framework.....	10
2.2 Nature of the applications.....	11
2.3 Public Sector Coverage .....	12
2.4 Regional and Administrative Heritage Distribution .....	12
2.5 Repeat assessments.....	14
2.6 Donor agency and partner government participation .....	16
2.7 Timeline and Publication.....	17
<b>Chapter 3 Evaluation of the Quality of PEFA Assessments</b> .....	<b>19</b>
3.1 Coverage of the Secretariat’s Quality Reviews .....	19
3.2 Review of Concept Notes / Terms of Reference .....	20
3.3 Review of Performance Reports - Introduction.....	21
3.4 Compliance in the Use of the Indicator Set.....	21
3.5 Incidence of No Scores.....	24
3.6 Adherence to the other sections of the PFM-PR .....	26
3.7 Repeat Assessments.....	28
3.8 Conclusions regarding the Secretariat’s Quality Reviews.....	29
<b>Chapter 4 Survey of Costs and Resource Use for Assessments</b> .....	<b>31</b>
<b>Chapter 5 Conclusions and Recommendations</b> .....	<b>33</b>
<b>Annex A PEFA Assessments used for assessing quality</b> .....	<b>37</b>
<b>Annex B Survey of the Costs of Implementing PEFA Assessments</b> .....	<b>41</b>
B.1 Introduction .....	41
B.2 Context and Approach.....	41
B.3 The Overall Cost of a PEFA Assessment.....	42
B.4 Co-financing arrangements.....	44
B.5 Composition of an Assessment Team .....	45
B.6 Explanatory factors.....	47
B.7 Conclusion.....	49

## List of Abbreviations

AfDB	African Development Bank
AsDB	Asian Development Bank
AusAid	Australian Agency for International Development
CFAA	Country Financial Accountability Assessments
CG	Central Government
CI	Compliance Index
CIFA	Country Integrated Fiduciary Assessment
CN	Concept Note
DFID	UK Department for International Development
DP	Development Partners
D-1, 2 or 3	Donor Practice Indicators
EC	European Commission
ERPFM	External Review of Public Financial Management
Framework	Public Financial Management Performance Measurement Framework
FY	Fiscal Year
IADB	Inter-American Development Bank
IMF	International Monetary Fund
MR	Monitoring Report
Norad	Norwegian Agency for Development Cooperation
OECD-BIS	Organization for Economic Co-operation and Development – Baseline
Indicator Set for procurement	
PEFA	Public Expenditure and Financial Accountability
PEMFAR	Public Expenditure Management and Financial Accountability Review
PER	Public Expenditure Review
PFM	Public Financial Management
PFM-PR	Public Financial Management – Performance Report
PFM-PR-SN	Public Financial Management – Performance Report – Sub-National
PI	Performance Indicator
SECO	Switzerland's State Secretariat for Economic Affairs
SNG	Sub-National government
TOR	Terms of Reference
WB	World Bank



## Executive Summary

This report represents the third study monitoring the roll-out of PEFA Framework application and the compliance with the methodology and principles set out in the Framework. It covers roll-out until October 2009 whereas monitoring of assessment quality and costs primarily includes the assessment reports received and reviewed during the period April 2007 to March 2009.

### Conclusions

#### *Roll-out*

- The number of completed PEFA assessments has continued unabated at a rate of 35-40 assessments p.a, reaching 151 substantially completed assessments by October 2009, covering 102 countries.
- A recent drop in recorded ongoing work combined with a drop in concept notes/terms of reference sent to the Secretariat for review has been noted. Further monitoring and investigation will be undertaken to establish the causes.
- PEFA assessment reports are increasingly in the format of a stand-alone PFM-PR.
- An increasing share of assessments covers a sub-national government entity.
- Country coverage of baseline assessments is reaching saturation level in Sub-Saharan Africa (only five countries not covered) whilst other regions are in the 50-70% range, excluding work planned but not commenced – with the exception of Western Europe and North America.
- The World Bank and the European Commission continue to dominate as the lead agencies, together being in charge of 85% of the assessment work.
- Repeat assessments are emerging in significant numbers, but some represent an attempt to create a more generally accepted baseline than the first assessment could muster and do not track performance changes from the earlier assessment.
- As only 3-4 years have lapsed since the first assessments were completed, it is not surprising that most assessments to date have not followed the recommended 3-5 years interval between baseline and repeat assessments. However, short intervals combined with frequent shifts in the leading agency could indicate that assessments are not done for commonly agreed purposes in a well-coordinated manner.
- An increasing share of finalized reports is becoming available to the public through the internet (up from 42% in 2007 to 56% in October 2009) though substantial scope for improvement exists.

#### *Compliance*

- The number of assessments being subject to the Secretariat's quality reviews has increased by 48% in FY09 compared to previous years and indicates close to total coverage. However, review of concept notes/ terms of reference remains at a low –

possibly decreasing – level, corresponding to only a quarter of the assessment roll-out rate.

- At the planning stage (concept note/terms of reference) the main issues identified by the reviews include the role of the government in the assessment, the blending of the standard purpose of the assessment (common pool of information) with a more specific donor-related purpose, inadequate provision or planning of the resources required and unclear institutional coverage of the assessment. A substantial portion of these issues remain unresolved or undocumented at the draft report stage.
- The overall compliance with methodology in rating the 31 performance indicators is steadily improving for finalized reports, reaching 91% for the final reports received in the first nine months of FY09.
- Low compliance remains an issue for a few selected indicators such as PI-7, 15, 19, 27 and D-1. The same indicators (except PI-27) are also the most affected by ‘no scores’ due to lack of data.
- Exclusion of selected indicators from assessment without adequate justification is becoming less of an issue, though still a concern as regards the donor practice indicators.
- Gradual improvement is noted in the quality of ‘Summary Assessment’ sections, whereas a comprehensive description of the structure of the public sector remains an important quality concern as it often affects the clarity of the scope of the assessment and the relative importance of individual indicators.
- Repeat assessments appear to provide a good basis for tracking progress over time in just over half the cases analyzed. Frequent repetition of PEFA assessments and lacking attempts to track changes in performance is a particular concern in a few countries.

#### *Costs of implementing an assessment*

- The overall cost of a PEFA assessment is on average in the order of USD 126,000, but wide a wide range from USD 25,000 to USD 280,000.
- The number of labor days used is in average about 92 with a range of 30-275, but typically in the range of 75-100 days. These figures generally cover the assessment team only, and not the time spent by other government officials and donor staff.
- The assessment costs in USD are very similar for the World Bank and the EC, whereas the assessments led by bilateral donor agencies have been somewhat lower. However, measured in labor day inputs EC and bilateral agency assessments are very similar whereas the World Bank has used about 50% more labor days per assessment.
- The core assessment team typically consists of 3 persons (5 for the World Bank).
- Average cost per labor day is USD 1100-1300 per day for assessors funded by EC and bilateral agencies, but only USD 767 per day for the World Bank, partly reflecting differences in the way assessment teams are mobilized and contracted.

- Size of country – in terms of population - showed a significant correlation with the number of total labor days used to complete an assessment.
- It has not been possible to find any correlation between level of resource inputs and assessment report quality as measured by the Secretariat's compliance index.

## **Recommendations**

- Proactive government involvement in the assessment process should continue to be promoted, including enhanced training effort for government officials to play a key role in assessment implementation.
- The standing recommendation of undertaking formal and full repeat assessments every 3-5 years, and certainly not annually, remains valid.
- Partners will encourage task teams to share concept notes at the draft stage with the Secretariat for comment.
- PEFA assessment reports should disclose, as a standard, a statement on resource use in implementing the assessment and names of the assessment team members.
- The peer review process should be identified at the CN stage; quality assurance arrangement should be transparent, set out in the CN/TOR and explained in the full report.
- For reports on which the PEFA Secretariat has provided comments, Partners are encouraged to share with the Secretariat the revised versions so that the Secretariat may follow-up on how the comments were addressed.
- The Secretariat will develop guidelines for repeat assessments. The guidelines would include the recommendation to specifically set out in the CN/TOR the need for the assessment to track performance change since a specific earlier assessment.
- Repeat assessment teams should be provided Secretariat comments on the earlier finalized assessment report in order improve the basis for tracking progress.
- To strengthen Summary Assessments - to focus more on the relative importance of weak linkages – guidance and training on formulating this section should be enhanced.
- PEFA training should be strengthened to include a module on the structure of the public sector and its importance for distinguishing national and sub-national level in assessments.
- The feasibility of developing a standard compliance index for CN/TORs and monitoring of its development as is currently the case for the compliance index for indicator assessment should be investigated.
- As a supplement to the compliance index for indicator assessment, a standard method of monitoring compliance of other parts of the PFM-PR should be development.



# Chapter 1

## Introduction

This is the third monitoring report prepared by the Secretariat. It provides roll-out information up to October 2009 and analyses trends in roll-out of application of the Framework since the previous Monitoring Report 2007 (MR07). As regards analysis of compliance issues and a survey of resource use and cost of implementing PEFA assessments, it covers the period from April 1, 2007 to March 31, 2009. It assesses the quality of 71 assessment reports submitted to the Secretariat during the reporting period.

Chapter 2 provides an overview of the application of the Framework. Chapter 3 evaluates the quality of PEFA assessments reviewed by the Secretariat. Chapter 4 summarizes the finding of a survey of the costs of implementing PEFA assessments (with full details in Annex 2) and Chapter 5 contains conclusions and recommendations.

## Chapter 2

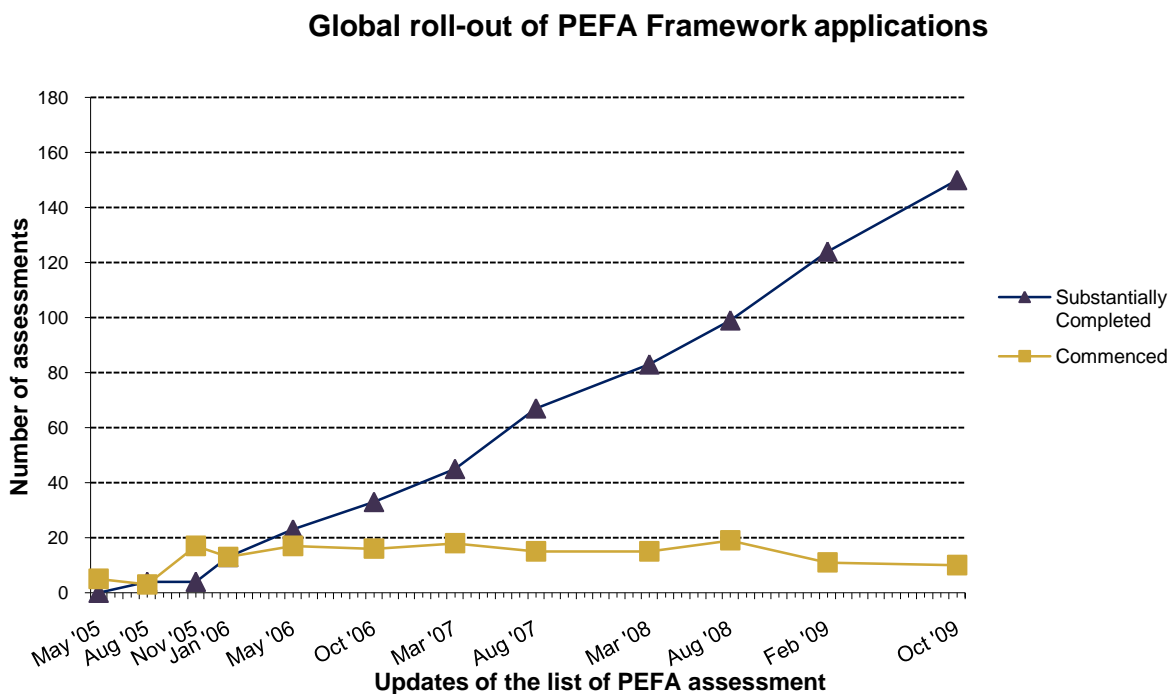
### Overview of Application of the Framework

#### 2.1 The rate of roll-out of the Framework

Diagram 1 shows the roll-out to beginning of October 2009. During the period of 52 months from the launch of the Framework in mid-June 2005 to October 9, 2009, a total of 151 PEFA assessments have been received by the Secretariat, 101 of which between April 1, 2007 and October 7, 2009. On average, this represents just over 3 assessments a month. An unusual drop in ongoing work is noted since February 2009. It is not clear if this is temporary, reflects a data capture problem, is associated with faster completion of reports or indicates a permanent change in the roll-out rate. Such a drop may spill over into the trend in completed reports only 6-12 months later. Up till October 2009, the trend in completed reports has continued to be on a steadily increasing trend.

The 151 assessments implemented cover 105 countries; the difference representing sub-national government (SNG) and repeat assessments. Included are assessment reports of Norway (prepared by Norad) and of Canton of Lucerne (prepared by a Swiss university).

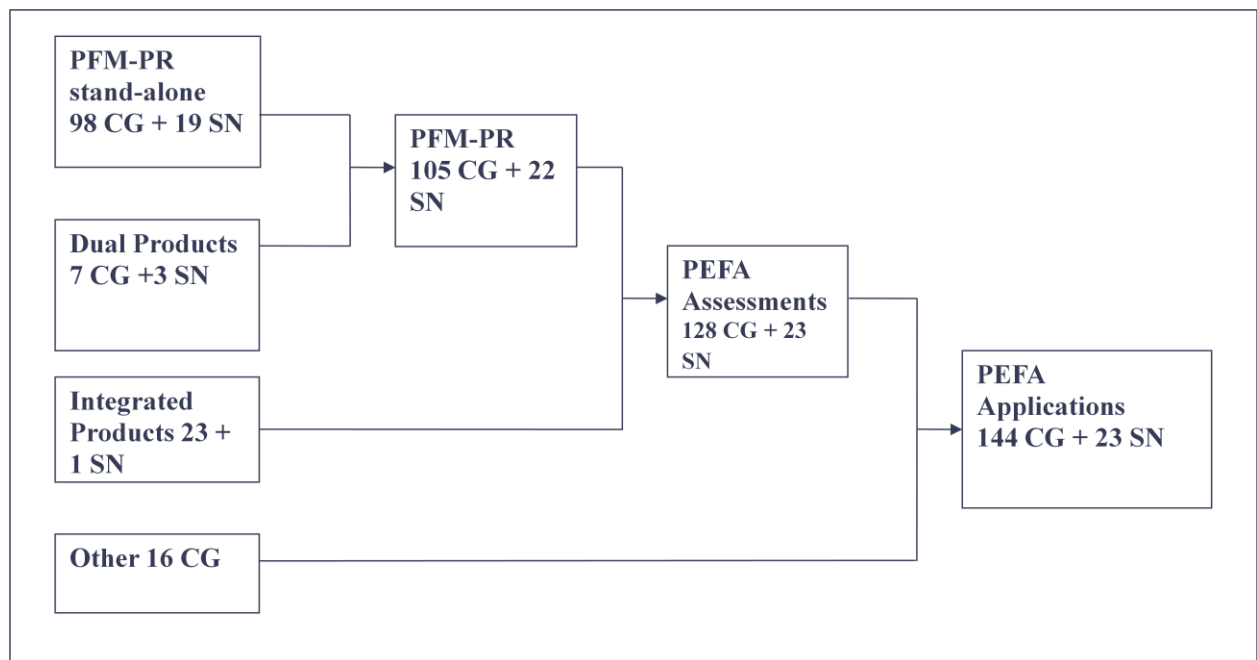
**Diagram 1: Global roll-out of the PEFA Framework as of October 7, 2009**



## 2.2 Nature of the applications

Diagram 2 shows the type of PEFA assessments received by the Secretariat in terms of: (i) Stand-alone PFM-PR; (ii) Dual product – part of a wider document, such as PEMFAR, PER, ERPFM, CIFA; and (iii) Integrated – section 3 of the PFM-PR integrated into a different analytical product, such as a CFAA.

**Diagram 2: Variety of PEFA Applications as at October 7, 2009**



For completeness, a number of other applications of the Framework is included in the diagram. These are not considered genuine PEFA assessments due to substantial deviations between the content of the Framework and the way in which it was adopted for the assessment. They represent PFM assessments, which used only a limited range of the Framework's performance indicators or did not use the scoring methodology.

In considering the variation from 2007 to 2009 in the spread across the various methods of application, there has been an increased use of the PFM-PR stand-alone at the central government (CG) and sub-national (SN) level, 20% and 33% respectively, a slight decrease in the use dual products at the CG level (-8%), and a larger decrease in the use of dual products at the SN level (-37%), and a slight decrease in the use of integrated products (-7%) and other CG applications (-5%).

## 2.3 Public Sector Coverage

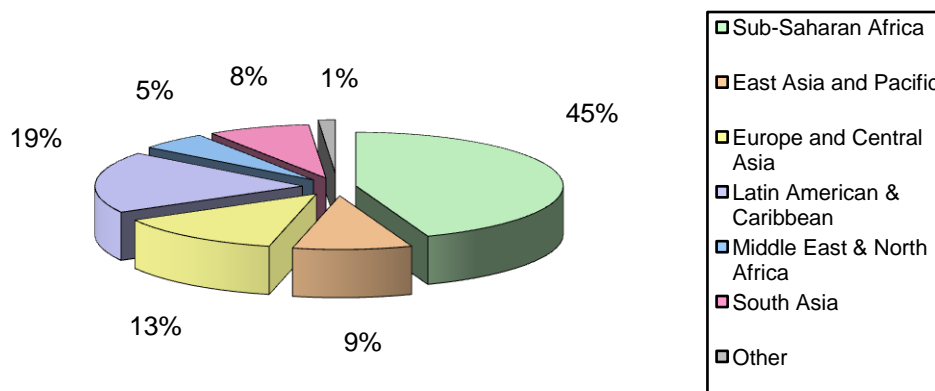
The vast majority of PEFA assessments continue to be at the central government level. Twenty-three SNG assessments were conducted during the reporting period (excluding the ones that deviated significantly from the Framework) out of the 71 assessments reviewed. Nine of these used the draft guidelines prepared by the Secretariat for the use of the Framework at SNG level, which is nearly a 70% usage rate if the 10 subnational reports conducted during or before the month the draft subnational guidelines (March 2008) were issued are not considered.

An attempt of using the PEFA Framework for a PFM assessment for the health sector was conducted for Mozambique, the first of its type. This was received by the Secretariat in April 2009, but has not been considered in this report.

## 2.4 Regional and Administrative Heritage Distribution

The distribution of substantially completed PEFA assessments to date by region is shown in diagram 4 below. As at October 7, 2009, 45 percent of the reports have been conducted in Sub-Saharan Africa, over one-third of which relate to francophone Africa. The Latin America and Caribbean region is represented by 28 assessments (19%), of which 12 are for the Caribbean countries. East Asia/Pacific region is represented by 14 reports (9%). Europe and Central Asia is represented by 20 (13%) reports, Middle East and North Africa countries are by 8 (5%) reports and South Asia by 12 (8%) reports, most of which are at SNG level (India and Pakistan). This distribution is significantly influenced by the number of countries covered by each region. While there are no significant changes in distribution amongst regions since the last update, an additional category, titled “Other”<sup>1</sup> was added to account for the recent roll-out of the PEFA tool in countries which do not fit the regional classification used previously.

**Diagram 3: Regional Distribution of PEFA Application as of October 7, 2009**



<sup>1</sup> Other” includes Western European countries, North America, Australia, Japan and New Zealand.

Diagram 4 shows the country coverage by region. At 90% the coverage of Sub-Saharan Africa is reaching saturation levels for *baseline* assessments, subject to those cases where the baseline has not been finalized due to lacking support or acceptance of the draft by the government. Only five countries in this region have not been covered by a baseline assessment (Eritrea, Somalia, Zimbabwe, Angola and Equatorial Guinea).

**Diagram 4: PEFA Assessment Coverage by Region (%)**

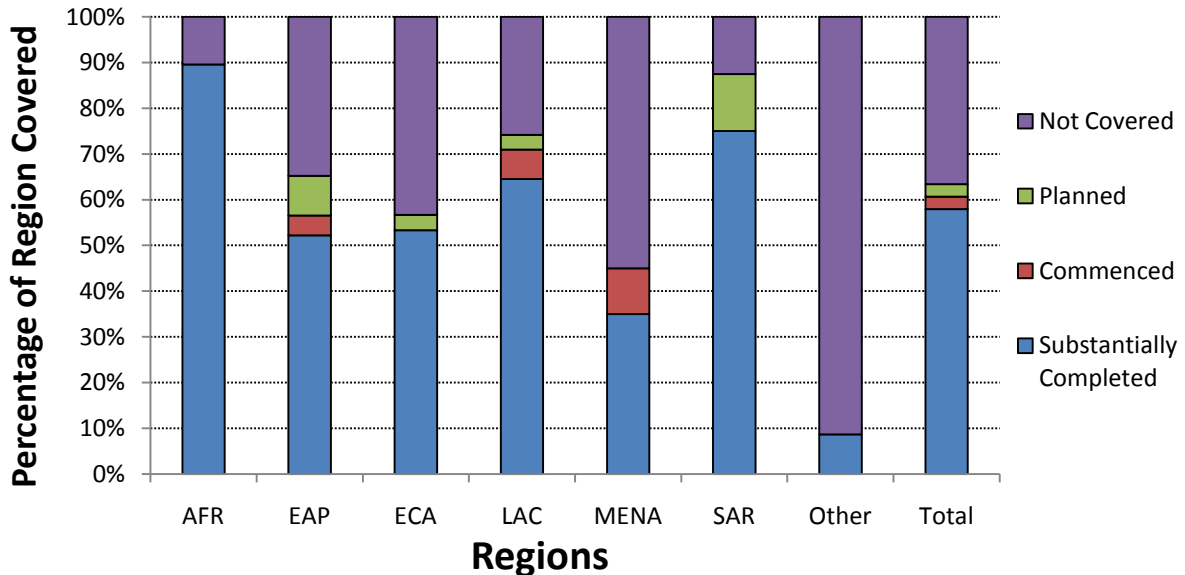


Table 1 shows the country coverage to date of assessments by region and administrative heritage. There is a significant increase in the number of assessments carried out in countries with a Russian administrative heritage, all from the Europe and Central Asia region, and in Sub-Saharan Africa with a British administrative heritage. The only relative decrease of the roll-out (in percentage terms) is amongst countries with a French administrative heritage.

**Table 1: PEFA Tool use by Region and Administrative Heritage as of October 7, 2009**

Number of countries	Origin of Administrative Heritage					
	Total	British	French	Spanish/ Portuguese	Russian	Other/Mixed
Region						
Sub-Saharan Africa	67	34	26	4		3
East Asia and Pacific	14	5		1		8
Europe and Central Asia	20				19	1
Latin American & Caribbean	28	12	1	15		
Middle East & North Africa	8	3	1			4
South Asia	12	9				3
Other	2					2
Total	151	63	28	20	19	21

## 2.5 Repeat assessments

As of October 2009, 26 repeat assessments have been substantially completed (i.e. to the full draft stage), of which 25 used the full Framework for both the first and subsequent assessments. The repeat assessments covered 21 countries i.e. 5 assessments represented a third or fourth assessment in a country. Table 2.a summarizes the situation. Four countries have had three assessments (Ghana, Malawi, Uganda and Tanzania), but only Ghana and Malawi can be counted as genuine repeat assessments; and Tanzania now has four. The situation concerning the third Tanzanian assessment (2007, updated in 2008 and still not finalized) is not at all clear, as it assesses progress in PFM performance in relation to the 2004 assessment (which used a different ‘draft’ version of the PEFA Framework) and does not reference the 2005 or 2006 assessments.

The Ugandan situation is also unusual. The Auditor General conducted a PEFA assessment in 2008, the report mainly focusing on Section 3 of the Framework (the section covering the scoring of indicators). This was published without any external quality review and did not reference the 2005 assessment for tracking of performance change. Though compliance with the Framework appeared satisfactory (if considered as a draft for review), a donor-led PEFA assessment was nevertheless conducted later in the same year.

As more repeat assessments are completed a more thorough analysis of them will be possible. However, in considering these 26 repeat assessments, there are a couple observations worth mentioning.

- Firstly, only 4 repeat assessments were completed within the time period suggested in the PEFA Framework (3-5 years). All other repeat assessments were completed

before the minimum recommended number of years. The Guinea Bissau assessment falls a month short of 3 years, but is included in this figure.

- Secondly, only 42% of the repeat assessments were completed by the same lead agency as the previous assessment.

**Table 2.a Countries with repeat PEFA Performance Assessments**

Country	1st Assessment		2nd Assessment		3rd Assessment		4th Assessment	
Afghanistan	Lead agency: WB Date: December 2005	Full PEFA assessment with PFM-PR	Lead agency: WB Date: June 2008	Full PEFA assessment with PFM-PR				
Bolivia	Lead agency: Self Date: October 2007	Full PEFA assessment with PFM-PR	Lead agency: WB Date: August 2009	Full PEFA assessment with PFM-PR				
Ghana	Lead agency: DFID Date: September 2005	Desk study with 20 indicators scored	Lead agency: WB Date: June 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: September 2009	Full PEFA assessment with PFM-PR		
Guinea Bissau	Lead agency: WB Date: June 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: May 2009	Full PEFA assessment with PFM-PR				
Honduras	Lead agency: WB Date: February 2006	CFAA update with full PEFA assessment with PFM-PR	Lead agency: EC Date: December 2008	Full PEFA assessment with PFM-PR				
Kenya	Lead agency: DFID Date: July 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: October 2008	Full PEFA assessment with PFM-PR				
Kosovo	Lead agency: WB Date: March 2007	Full PEFA assessment with PFM-PR	Lead agency: WB Date: March 2009	Full PEFA assessment with PFM-PR				
Kyrgyz Republic	Lead agency: DFID Date: January 2006	Full PEFA assessment with PFM-PR	Lead agency: SECO Date: August 2009	Full PEFA assessment with PFM-PR				
Lesotho	Lead agency: WB Date: June 2007	Full PEFA assessment with PFM-PR	Lead agency: DFID Date: July 2009	Full PEFA assessment with PFM-PR				
Madagascar	Lead agency: EC Date: May 2006	Assessment based on draft 2004 PEFA Framework	Lead agency: WB Date: May 2008	Full PEFA assessment with PFM-PR				
Malawi	Lead agency: EC Date: July 2005	Full PEFA assessment with PFM-PR	Lead agency: EC Date: August 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: June 2008	Full PEFA assessment with PFM-PR		
Moldova	Lead agency: EC Date: April 2006	Full PEFA assessment with PFM-PR	Lead agency: WB Date: June 2008	Full PEFA assessment with PFM-PR				
Mozambique	Lead agency: EC Date: March 2006	Full PEFA assessment with PFM-PR	Lead agency: Norway Date: February 2008	Full PEFA assessment with PFM-PR				
Papua New Guinea	Lead agency: WB Date: February 2006	Full PEFA assessment with PFM-PR	Lead agency: WB Date: October 2008	Updated rating table with little or no evidence				
St. Kitts and Nevis	Lead agency: EC Date: April 2007	Full PEFA assessment with PFM-PR	Lead agency: EC Date: September 2009	Full PEFA assessment with PFM-PR				
Tanzania	Lead agency: WB Date: May 2006	PEFA assessment as part of a PEFAR	Lead agency: WB Date: June 2006	Update of assessment table with substantial explanation but not a full report	Lead agency: WB Date: June 2008	Full PEFA assessment with PFM-PR	Lead agency: WB Date: March 2009	Full PEFA assessment with PFM-PR
Togo	Lead agency: WB Date: June 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: November 2008	Full PEFA assessment with PFM-PR				
Trinidad and Tobago	Lead agency: EC Date: June 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: October 2008	Full PEFA assessment with PFM-PR				
Uganda	Lead agency: EC Date: May 2006	Full PEFA assessment with PFM-PR	Lead agency: Auditor General of Uganda Date: March 2008	Full PEFA assessment with PFM-PR	Lead agency: WB Date: December 2008	Full PEFA assessment with PFM-PR		
Vanuatu	Lead agency: EC Date: July 2006	Full PEFA assessment with PFM-PR	Lead agency: EC Date: August 2009	Full PEFA assessment with PFM-PR				
Zambia	Lead agency: DFID Date: December 2005	Full PEFA assessment with PFM-PR	Lead agency: Self assessment Date: June 2008	Full PEFA assessment with PFM-PR				

**Table 2.b Countries assessed in FY06 for which no repeat assessment implemented**

<b>Applications</b>	<b>Date</b>	<b>Region</b>	<b>Lead agency</b>	<b>Type of assessment</b>	<b>Status</b>	<b>Repeat planned</b>
Congo Brazzaville	Mar-06	AFR	EC	Dual	Finalized	No
Tanzania-Local govt.	May-06	AFR	WB	PFM-PR-SN	Finalized	No
Fiji Islands	Jun-05	EAP	WB	Dual	Finalized	No
Lao PDR	May-06	EAP	EC	Integrated	Finalized	2009
Syria	Mar-06	MENA	IMF	Integrated	Finalized	No
Bangladesh	Apr-06	SAR	WB	Integrated	Finalized	No

Table 2.b lists the six cases where an initial PEFA assessment took place in the first year after the launch of the PEFA Framework (i.e. in FY06) but where a repeat assessment has not yet been undertaken. In none of these six cases did the initial assessment represent a standard PFM Performance Report. In one case, a repeat had been planned. Overall, of the 23 countries with an initial assessment in 2006, 74% have had a repeat assessment within a period of three years. The quality of repeat assessments is discussed in Section 3.7.

## **2.6 Donor agency and partner government participation**

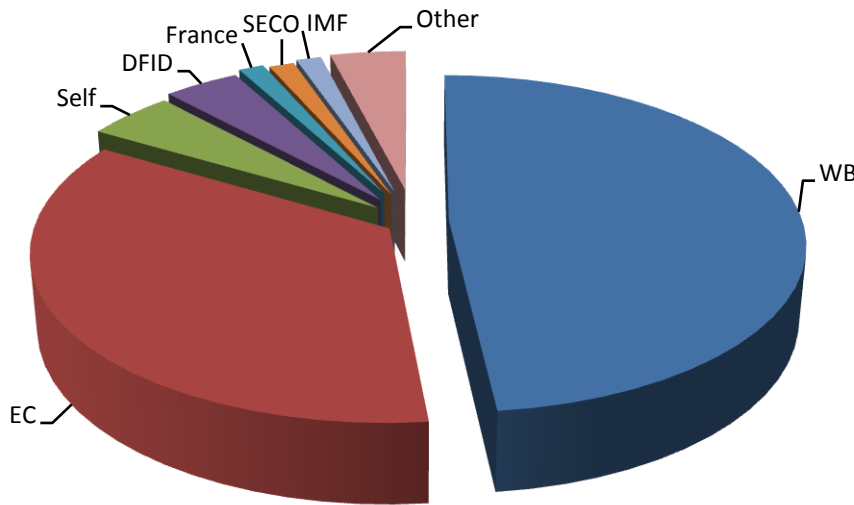
Most assessments have been initiated and led by donor agencies, but the extent of government involvement has increased sharply over the last 2 years. Self-assessments (conducted by governments themselves, using either their own staff or non-government organizations contracted by them) and joint assessments are still a very small proportion of assessments, but the Secretariat expects that this proportion will rise over the coming years. Governments are becoming increasingly aware of the benefits of PEFA assessments, in terms of identifying weaknesses in PFM system performance and therefore providing a baseline that can inform the strengthening of PFM reform strategies. In fact the categorizing of PEFA assessments in terms of donor-led, joint and self has become increasingly blurred, as the involvement of governments increases.

The World Bank and EC continue to dominate in terms of leading agencies in PEFA assessments; the World Bank has been the leading agency in 48% of assessments, the EC 35%. Diagram 5 summarizes the situation. The “Other” category represents assessments that are led by the donor institutions that have only led one assessment and are not PEFA partners.

World Bank-led assessments are often part of “integrated” or “dual purpose” assessments (usually the latter). Other agencies that have played a leading role are DFID, France, Switzerland, IMF, AsDB, IADB, AfDB, AusAid, and Norway. Donor agencies that are not the lead agencies are often involved in the assessment through funding or contributing to the quality assurance process (usually through donor PFM groups formed in the assessed country).



**Diagram 5: Lead agencies of PEFA Assessments**



## 2.7 Timeline and Publication

**One hundred and two** PEFA assessments had been finalized as of October 7, 2009. Forty-nine assessments have yet to be finalized. An issue, as highlighted in the previous MR07, is the significant number of reports that seem to take a long time to finalize. This is the case for “dual purpose” reports (typically led by the World Bank) and to a lesser extent stand-alone reports (typically led by EC). Finalizing of the dual purpose reports can take over a year. For example, the Secretariat commented on the Sindh province (Pakistan) and Himachal Pradesh (India) PEFA assessment reports in February 2008. The Himachal Pradesh report was finalized in January 2009 and the Sindh assessment finalized in February 2009.<sup>2</sup>

Of the finalized reports, 57 have been published, and therefore appear on the PEFA website. Table 4 summarizes finalization and publication of PEFA assessment reports

<sup>2</sup> According to information provided in January by the Bank’s resident office in Pakistan, the report would be finalized in March.

**Table 3: Finalization and Publication of PEFA Assessment Reports, October 2009**

<b>Lead Donor</b>	<b>Number Led</b>	<b>Number Final</b>	<b>Number Public</b>	<b>Final</b>	<b>Public out of Final</b>
World Bank	73	44	29	60%	66%
EC	53	41	21	77%	51%
Other	25	17	7	68%	41%
Total	151	102	57	68%	56%

Since the status reported in the MR07 the percentage of final reports made available to the public through the internet has increased from 42% to 56%. This is an encouraging achievement, though there is still scope for improvement. The share of finalized reports to total number of report commenced remains almost constant overall, but with a diminishing difference between the lead agencies.

## Chapter 3

### Evaluation of the Quality of PEFA Assessments

#### 3.1 Coverage of the Secretariat's Quality Reviews

Data on formal quality reviews from the Secretariat's annual progress reports FY07, FY08 and FY09 are shown in table 7. For FY09 the data shows a large increase in the number of assessments on which the secretariat has been requested to undertake a quality review, from 29 to 43. This is a new development as the change from FY07 to FY08 was quite marginal. At the same time, the semi-annual stock-taking of PEFA assessments ongoing and reaching substantial completion reveal an extremely stable trend of about 35-40 assessments reaching substantial completion annually and some 15 assessments ongoing at any time.

**Table 4 Number of quality reviews undertaken by the Secretariat on Request**

Operational year	Reviews of CN/TOR	Initial report reviews	Follow-up reviews
FY07	Not reported	27	Not reported
FY08	13	29	Not reported
FY09	10	43	7

Comparing those trends could lead to two conclusions:

- The coverage of report reviews by the Secretariat is approaching 100% (although there are still reports known to be completed without secretariat review). This is good news.
- There are PEFA based assessments which are not captured by the semi-annual stock-taking, which later turn up for report review. This has been noted e.g. for state level assessments in Nigeria. This suggests an information gap to be closed in the future.
- A less encouraging trend, however, is the decline in the number of concept notes and terms of reference reaching the Secretariat for a formal review. The number corresponds to only a quarter of the assessments being undertaken and on a declining trend.

This is not to say that the Secretariat has not been consulted on planning of assessments, but most of such involvement has been in terms of giving advice on specific issues the planners may have decided to seek guidance on. It does not provide the Secretariat with an opportunity comment on the reflection of good practices in the process of planning assessments, coordination of PFM analytical work in general and full involvement of stakeholders. In many cases, the terms of reference for PEFA assessments are obtained by the Secretariat through informal contacts with the private consulting sector rather than from the PEFA partners and other lead agencies involved.

## 3.2 Review of Concept Notes / Terms of Reference

For the analysis in this monitoring report Concept Notes/Terms of Reference were examined for 20 assessments<sup>3</sup> where the Secretariat had been requested to comment during the period covered. Of the reviews performed by the Secretariat, eleven concerned concept notes (CN) prepared by the World Bank, six were for terms of reference (TOR) prepared by the European Commission and three were CN/TORs prepared by bilateral donors.

The analysis here only relates to comments that highlight deficiencies in the CN/TOR. The comments were classified over 20 areas (background information, purpose, scope, etc) drawn from the TOR checklist issued by the Secretariat. Each of these areas are important either for the quality of the assessment and /or for the assessment's role as a common information pool within the Strengthened Approach to Support PFM Reform.

Revised CNs/TORs were compared with the original drafts for evidence of response to the comments. Responses were classified as full, partial or none/unsatisfactory. In a few cases, no revised CNs/TORs were submitted to the Secretariat, and the Performance Reports were examined for evidence of response. In some cases, it was not possible to judge whether the comments had been accepted.

For the 20 assessments, a total of 97 comments were made, with an average of 5 areas commented on for each CN/TOR out of the 20 areas analyzed. Almost all areas were commented on, but comments were most frequently made on:

- the role/involvement of the government (60% of reviews),
- the purpose of the assessment (45%), typically on intentions to incorporate a donor fiduciary instrument or a reform action plan in the PFM-PR.
- the composition of the team and budgeted inputs (45%)
- coverage of the assessment (40%), typically regarding the status of geographically decentralized, government entities.

Arrangements for external validation were commented on in 25% of the reviews, often because the PEFA Secretariat was not mentioned as a reviewer or it was listed as the only external reviewer.

Responses were tracked on 82 of the comments. Of these, 45 (55%) were fully addressed, 10 (12%) partially addressed, and 27 (33%) were not responded to or unsatisfactorily answered, the latter in some cases due to misunderstanding of the comment.

Comments on the role of government and arrangements for external validation were reasonably well addressed (fully addressed in 50% and 80% of cases respectively), whereas comments on the other three high frequency areas of comments were generally not satisfactorily addressed (at about 33% fully addressed).

---

<sup>3</sup> Due to IT problems, three of the 23 reviews done could not be traced

### 3.3 Review of Performance Reports - Introduction

The review of performance reports in this monitoring report covers 71 countries/sub-national governments for which PEFA assessments were reviewed by the Secretariat between April 1, 2007 and March 31, 2009. Forty four of these are now in revised/final draft stage. The Secretariat had not received revised/final drafts for the remaining 27 at the time, either because the assessments were relatively recent (seven since January 1, 2009), or because the revised/final drafts have not yet been submitted to the Secretariat. The 44 revised/final drafts include eight carried over from the previous reporting period (up to 31 March 2007), reflecting initial drafts reviewed by the Secretariat during that period. Out of the 71 assessments, 50 were baseline assessments, 12 were repeat assessments and 9 covered sub-national governments (SNG).

The report does not cover reports that were not submitted to the Secretariat for review, but were received by the Secretariat as final drafts: the Nigerian state of Anambra, Nepal, and Democratic Republic of Congo. A report for Russian Federation was received, but a review not requested. The report also does not cover assessment reports received that did not sufficiently comply with the Framework to be considered a fully PEFA-based assessment: Bolivia, Former Yugoslav Republic of Macedonia and one Indian state.

The rest of this chapter discusses the following aspects of the report reviews: Compliance with regard to use of the indicator set, the incidence of No Scores, coverage of the Secretariat's quality reviews, and a discussion of repeat assessments.

### 3.4 Compliance in the Use of the Indicator Set

The compliance in the use of the indicator set has increased steadily over time, as seen below in tables 5.a and 5.b and diagram 6 below.

**Table 5.a: Aggregate Compliance Index for Revised/Final Drafts<sup>4</sup>**

	Reviewed June 2005-March 2007	Reviewed April 2007- March 2009
Number of Reports	47	43
Average Compliance Index	59%	80%

The analysis of compliance indices shows two positive improvements from the previous reporting period:

<sup>4</sup> The number of such drafts reviewed in connection with the current reporting period includes 8 reports for which the initial draft was submitted to the Secretariat during the previous reporting period but which had not been substantially completed/finalized by the end of that reporting period.

- The CI for revised/final drafts increased sharply to 80% for the current reporting period from 59% for the two previous reporting periods combined (ref. table 5.a).<sup>5</sup>
- The difference between the CI for the revised/final draft and the initial draft was significantly larger (21 percentage points) than for the June 2005-March 2007 period (10 percentage points) covered by the previous monitoring report. This indicates an increase in the extent of response by assessment teams to peer reviewer comments, including those of the PEFA Secretariat (ref. table 6).

The progression in quality over time is better illustrated in table 5.b and the related diagram 6, which indicate the index year by year according to the year in which the final report was issued. The data indicates that compliance is approaching 100% for finalized reports.

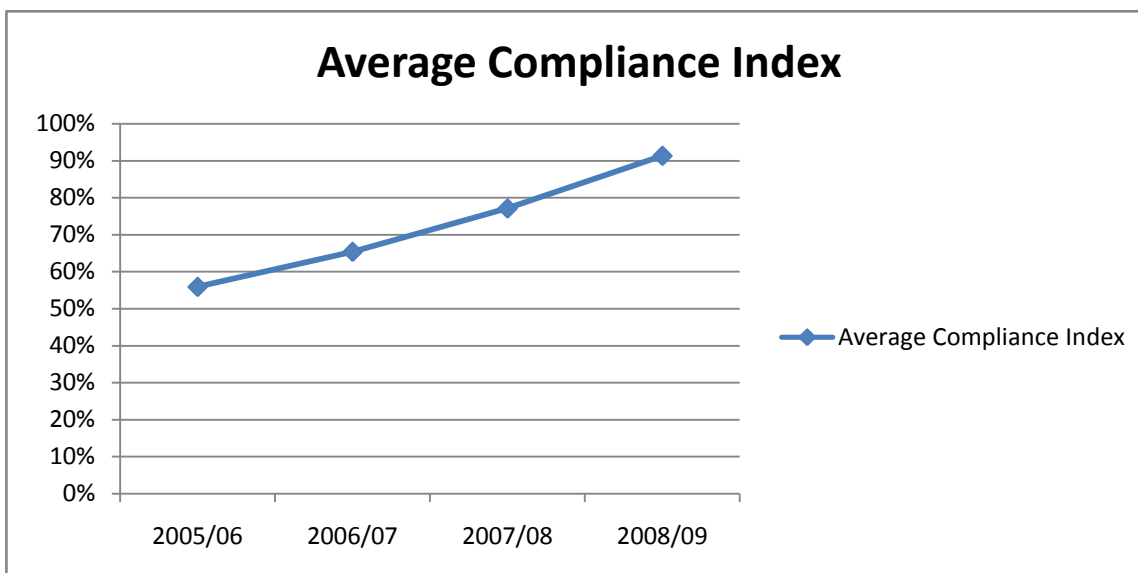
**Table 5.b Aggregate Compliance Index by Year of Finalization**

Fiscal Year Report was Completed <sup>1</sup>	2005/06	2006/07	2007/08	2008/09 <sup>2</sup>	Total
Number of Finalized Reports	19	24	23	6	72
Average Compliance Index	56%	65%	77%	91%	69%

<sup>1</sup> The World Bank Fiscal Year period is used, July – June.

<sup>2</sup> Only an eight month period is considered, July 2008 – February 2009 since February was the last month an official status of assessments was completed by the time this analysis was done.

**Diagram 6: Aggregate Compliance Index for Revised/Final Drafts**



<sup>5</sup> The previous monitoring report assessed compliance trends in relation to the first monitoring report. Compliance improved to 60% from 48% (page 28 of the 2007 Monitoring report). For ease of reading, this monitoring report assesses compliance trends in relation to the previous two reports combined.

A question remains, however, as to why the eight month period of FY09 (July-February) only saw six finalized reports as compared to four times as many in the each of the previous two years.

Table 6 shows the compliance index for the reporting period of the present report for groups of indicators represented by each core dimension plus the donor practices dimension. The second and third columns show the initial CI and revised/final CI for the 44 revised/final reports. The last column shows the initial CI for those 27 countries for which revised/final drafts have not yet been submitted to the Secretariat.

The compliance appears rather even across the core dimensions. The index for the first core dimension on budget credibility is generally higher than average, probably due to the quite straight forward quantitative nature of the four indicators involved. An index slightly lower than average is noted for the core dimension #4 on budget execution, especially for the report where a revised version was not subsequently received by the Secretariat.

**Table 6: Compliance Index by Core Dimension/Donor Practices**

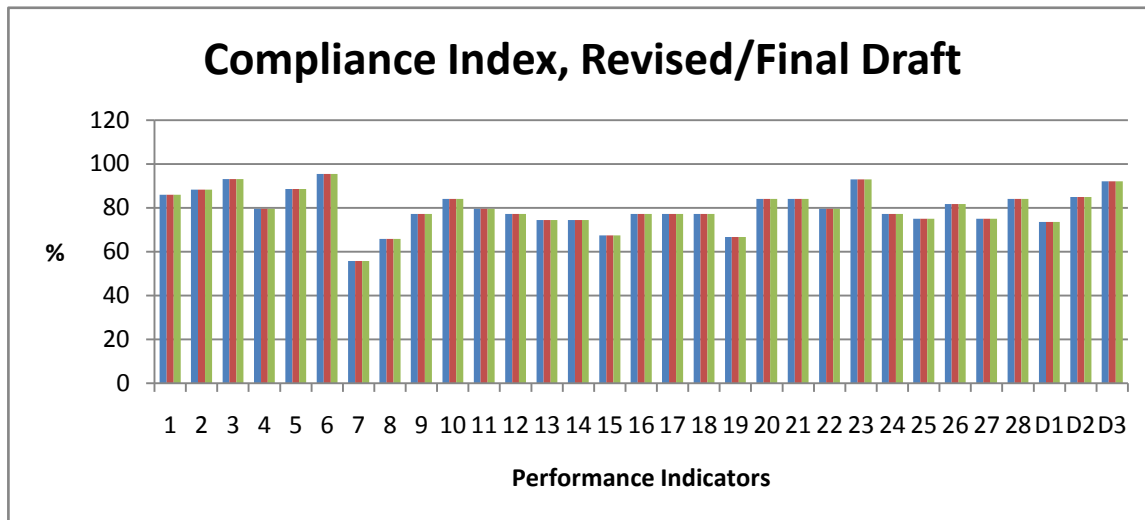
<b>Core Dimension</b>	<b>CI (%) Initial drafts 44 reports</b>	<b>CI (%) Revised/final drafts 44 reports</b>	<b>CI (%) Initial drafts 27 reports</b>
CD 1	70	87	63
CD 2	54	78	60
CD 3	59	78	50
CD 4	55	76	49
CD 5	61	81	62
CD 6	55	80	65
Donor Practices	65	84	59
<b>Overall CI 2007-2009</b>	<b>59</b>	<b>80</b>	<b>57</b>
<b>Overall CI 2005-2007 (46 reports)</b>	<b>48</b>	<b>58</b>	

The CI for the 27 initial draft reports, that had not been finalized, is about the same as the initial CI for the 44 reports now in revised/final draft stage: 57% versus 59%. If the same pattern holds, the final CI for these 27 drafts should also be significantly higher than the CI for the initial drafts.

Diagram 7 shows indicator scores for each indicator scored in the 44 revised/final drafts. The only indicator with a CI lower than 60 percent is PI-7 (56%). The previous monitoring report showed six indicators with low CIs (7, 8, 10, 15, 22, D-1), which,

moreover, were much lower, ranging from 29% to 43%. Performance indicators with CIs between 60% and 70% total three (PIs 8, 15, 19). Performance indicators with CIs between 70% and 80% total 11. Just over half (16) the indicators had CIs of 80% and over.

**Diagram 7 Compliance Index across indicators: Revised/Final Draft**



Comparison to the CI by indicator for the previous monitoring period (2005-2007) shows the changes in the CI between the initial and revised/final drafts. The largest changes are for PIs 4, 7, 15, 22, and 27.

Some specific observations based on a perusal of all internal evaluation reports prepared as inputs into this report:

- For some countries, there was no or very little change between the CI for the initial draft and for the final report, indicating the possibility that the Secretariat’s comments had been largely disregarded. These are: Haiti, Guinea, Sao Tome, Tuvalu;
- Seven reports went through two rounds of review by the Secretariat before emerging in final report status, thereby enhancing the final quality of the report: Moldova, Montenegro, Maharashtra, Uganda (though final draft had not been received at the time), Malawi, Seychelles, Liberia, and Afghanistan (through queries submitted by the team in response to Secretariat’s comments on the initial draft).

### 3.5 Incidence of No Scores

The Compliance Index is a numerical statistic that divides the number of correctly supported scores (based on sufficient evidence and correct score for the evidence) by the



numbers of indicators used. The number of indicators used may be less than the 31 Framework indicators for mainly three reasons:

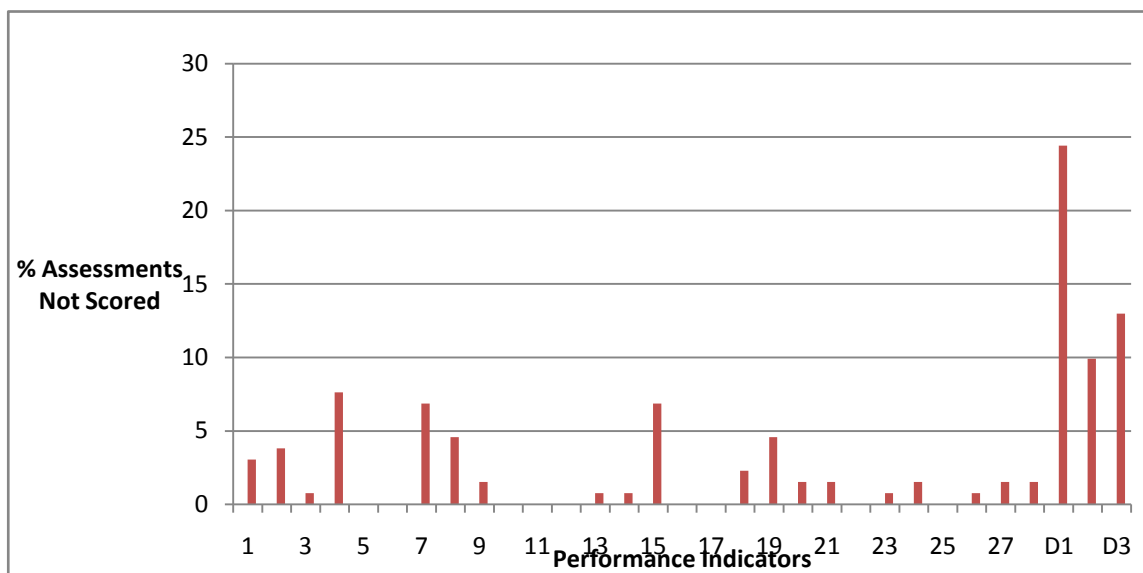
- Some indicators are not used because they are *not applicable* (NA) due to the country circumstances. This is most commonly the case where there are no SNGs (thus PI-8 is not applicable) and/or there is no donor budget support (D-1) or no (insignificant) donor project support (D-2-D-3). In this case, it is of course legitimate to measure the CI as the number of correctly supported scores by the number of indicators used (subtracting the NA indicators from the maximum 31 indicators).
- Some indicators are *not used* (NU), because of a deliberate decision taken prior to the assessment not to use them, even though they can potentially be scored. The CI is then measured in the same way as for NA indicators. Countries where indicators have deliberately not been used, even though they could have been used, include Kosovo (DP indicators), Tuvalu (DP and revenue administration indicators), Dakar city, Sao Tome (DP and external oversight and scrutiny indicators), Senegal (DP indicators), and the four Brazilian states (DP indicators).
- Some indicators are *not rated* (NR), because the assessment team was not able to obtain the information necessary to rate (due perhaps to meeting scheduling problems and/or non-cooperation from the authorities). The CI is then measured the same way as shown above. Frequently such cases relate to inadequate information to score just one of the dimensions of the indicator, whereas other dimensions would be correctly rated.

Diagram 8 shows the incidence of Not Scored indicators in terms of the percentage of indicators that have not been scored. Each bar represents the percentage of instances where an indicator was not scored.

The diagram indicates that that the DP indicators represent the highest incidence (also the case in the 2007 monitoring report). D-1 has 25 instances, but 16 of these are NA. Indicators D2-D3 have 22 instances, of which only 4 are NA, so the issue of deliberately not using these indicators is more serious. Encouragingly, the instances for D2-D3 have dropped from 30% and 35% in the 2007 Monitoring Report to 10% and 14% for D2 and D3 respectively. The other indicators with relatively high instances of no scores are PI 4, PI 7, PI 15 and PI-8. However, as indicated above, the no scores for the first three mainly reflect the wording of the scoring criteria. The incidence for PI-8 reflects its NA status for some reports.

Some indicators do not have any no scores for this current monitoring period, but did in the previous monitoring report: these are PIs 12, 16, 17, 22, 23 and 25.

**Diagram 8 Incidence of indicators not scored**



It should be noted that diagram 8 is based on all 71 countries covered by this monitoring report. With 27 of these still at the initial draft stage, it is possible that the incidence of NR may fall in the revised/final draft, if evidence is found that justifies a score.

The PEFA Framework is designed as an integral framework and all indicators should be used, unless they are clearly not applicable. In the latter case, justification for omitting the indicator should be provided in the report.

### **3.6 Adherence to the other sections of the PFM-PR**

Table 7 summarizes the instances where the Secretariat has recommended significant strengthening of sections other than those dealing with the performance indicators i.e. sections 1, 2, 4 and the Summary Assessment. The table also shows the extent to which the revised/final draft satisfactorily addressed those comments. The table shows that about 60 percent of the Secretariat’s recommendations are taken into account in the revised/final draft.

Table 7 refers to the 44 reports for which a subsequent version of a report was received. Strengthening measures were similarly recommended in most of the above categories for those countries where the revised/final draft has not yet been received.

**Table 7: Adherence to the other sections of the PFM-PR**

<b>Section</b>	<b>Initial Draft # Countries where strengthening recommended 1/</b>	<b>Revised/Final Draft # Countries where Recommendations addressed 2/</b>
Summary Assessment	27	17 (63%)
Introduction sect. 1	21	12 (57%)
Country Background sect. 2	24	14 (58%)
Government Reform Process section 4	16	10 (63%)
Country Specific Features section 3.8 3/	1	1 (100%)

1/ Covering only those countries where revised/final reports are available.

2/ Where recommendations are partially incorporated, countries are counted as 1/2

3/ Only 2 reports appeared to have a Country Specific issues sub-section (Azerbaijan and Solomon Islands).

The main findings are:

- The most frequent section needing strengthening is the Summary Assessment. The comment in the previous MR07 (paragraph 89) that the first part of the Summary Assessment tends to be a listing of indicators and the discussion on the implications for budgetary outcomes therefore suffers appears to have been mitigated to a large extent, though clearly further improvement is possible. More story lines are being prepared (as recommended in the previous report). The Summary Assessment is becoming a more useful tool for showing the way forward to strengthening a government's PFM reform strategy, though more could be done to prioritize the addressing of the weakest linkages, including making much greater use of the matrix in Appendix 1 of the Framework. Improvement in quality in this area may be due to training in the use of the Framework and experience gained through the conduct of multiple PEFA assessments.
- A common omission in the Introduction is a comprehensive description of the structure of the public sector in terms of number of entities, relative importance and fiscal relations e.g. in a table. This omission negatively affects the understanding of what parts of general government the assessment covers and the relative importance of some indicators, notably PI-8 and PI-9.
- As noted in Table 6, very few reports have contained a sub-section on Country-Specific issues. Such issues are usually described elsewhere in the report, typically in the Background section 2.
- Section 4 of a standard PFM-PR on Government Reform Process has been omitted from several reports.

### 3.7 Repeat Assessments

Repeat assessments had been substantially completed for 12 countries at end of March 2009<sup>6</sup>. The quality of these varies widely. The main purpose of repeat assessments is to track performance since the last assessment. In order to do so, it is necessary to make sure like is compared with like. This means checking for example, if any mistakes were made in the scoring methodology in the previous assessment, or whether different data was used at arriving at the score (perhaps because audited accounts for the years covered by the earlier report may have different figures for revenue and expenditure than those used by the assessors). It also means checking that the coverage and definitions of terms are the same for both assessments.

Table 8 summarizes the salient features of the repeat assessments conducted to date.

**Table 8: Repeat Assessments**

Country	Previous scoring issues identified?	Data change issues identified?	Definition issues identified?	Progress validly tracked?
Madagascar		9 PIs		Yes
Mozambique	1 PI	2 PIs		Mainly
Moldova	3 PIs	1 PIs		Mainly
Afghanistan		2 PIs	2 PIs	Uncertain
Malawi	6 PIs	1 PI	1 PI	Mainly
Uganda	6 PI	7 PI	1 PI	Mainly
<i>Zambia I/</i>	<i>3 PIs</i>			<i>Mainly</i>
<i>Kenya I/</i>	<i>1 PI</i>		<i>1 PI</i>	<i>Mainly</i>
<i>Togo I/</i>				<i>No</i>
<i>Tanzania I/</i>				<i>No</i>
<i>Trinidad I/</i>	<i>Yes, many</i>		<i>1 PI</i>	<i>Partly (for 4-5 PIs)</i>
<i>PNG I/</i>				<i>No</i>

I/ Final drafts not received at the end of March 2009.

The main issue is comparability when the CIs are low in either or both years, as is the case for some of the countries mentioned above. A satisfactory assessment is one that verifies the basis on which the scores from the previous assessment were assigned, taking note of potential mistakes and/or non-comparability with the current assessment, and assessing change of performance over time on that basis. For general comparability, compliance indices of both the earlier and repeat assessments should be reasonably high. Only after this verification process has been carried out is it valid to track progress in

<sup>6</sup> Compared to the 21 countries with repeat assessments listed in table 2.a above, seven assessments were received after March 2009, and two assessments were considered new baselines by the lead agency so no comparison was envisaged.

PFM performance based on changes in scores. To help the verification process, it would be very helpful for assessment teams to have access to the Secretariat comments on the final report of the first assessment (a problem has been that assessment teams may have a different view of the correctness of the score from the first assessment than that of the Secretariat) and to the previous assessors' detailed notes which may not be reflected in the report.

Only the Madagascar assessment appears to be almost completely validly tracked, whilst about half of the repeat assessments reviewed provide a fairly good basis for tracking progress. Five of the twelve repeat assessments provide a poor basis for tracking performance changes, either because no attempt at all was made to track performance change or because the compliance rate of one or both reports was too low to provide an adequate basis for comparison.

Some repeat assessments re-scored the indicators of previous assessments. This practice is dangerous as it means changing an assessment typically without consulting and agreeing such changes with the team responsible for the previous assessment. Subsequent assessment may then dispute those re-scores, potentially leading to endless disputes.

### **3.8 Conclusions regarding the Secretariat's Quality Reviews**

The steady improvement in the compliance index are encouraging signs which could reflect the impact of four years of consistent training and dissemination efforts, an increasing stock of experienced PEFA assessors, combined with a high coverage of report quality reviews especially since the beginning of FY09.

The compliance index, however, reflects only the provision of sufficient evidence and the correct use of that evidence to score the indicators. While this is useful for a number of purposes, it cannot verify the accuracy and completeness of the data/information provided as evidence, nor does it measure the quality of the remainder of the report. Accuracy and completeness of the data is contingent on all major stakeholders in a country participating in providing the data and reviewing that it has been properly used in the report. It is also contingent on adequate resources being available for the assessment work, especially the quality, composition, time and logistics resources of the assessment team. These are factors that are determined at the CN/TOR stage and hard to change later. They should - and are to varying degree - described in the draft report, and such a description - if well done - could provide assurance that the evidence obtained and used in comprehensive and accurate. It is usually too late, though, to suggest significant amendments at the report stage.

The program needs to consider the following quality assurance measures:

- Ensure that the CN/TOR for all planned assessments undergo a quality review by the Secretariat (and preferably another major stakeholder at country level)

- Development of a quality/compliance index for CN/TORs (using the Good Practice Note as a starting point for deciding standards) and monitoring of its development as is currently the case for the compliance index for indicator assessment.
- Development of a quality/compliance index for the other parts of a PFM-PR – as a supplement to the compliance index for indicator assessment - and monitoring of its development in conjunction with the compliance index for indicator assessment

## Chapter 4

### Survey of Costs and Resource Use for Assessments

To provide a reference point to assist managers at the resource allocation stage of preparing for a PEFA assessment, the cost information of a sample of 30 assessments (15 led by the EC, 10 by the WB, and 5 by other PEFA partner organizations) was collected and analyzed. The sample was not chosen randomly; all 71 assessments for this monitoring period were considered. The decision to include an assessment was based on the availability of the budget envelope or actual cost and its composition.

The analysis comes with a few caveats. First, the labor hours spent for peer review were not included because the majority of the reviewers did not code their time separate from other daily/project tasks. Second, the totals were converted to U.S. Dollars when in another currency, using the exchange rate of the first day of the month the main mission started. Third, countries are included that have some financial contributions missing but were considered insignificant. Fourth, due to the limited sample size, the financial figures used are a mixture of budget and actuals.

The average (mean) cost of a PEFA assessment was found to be \$ 126,000. It was similar across regions and across the EC (\$137,000) and WB (\$134,000), but significantly lower for those led by other PEFA partner organizations (\$75,000). Of the 5 assessments in this category, Guinea may be an anomaly, while the zero and low non-labor costs of Norway and Switzerland-Lucerne Canton respectively explain their lower overall costs. When the days of labor used (or planned) by the lead organization is considered, the average of the EC (79) and those led by other PEFA partner organizations (76) are close and the WB (123) is significantly higher, however, a smaller sample size due to insufficient information for four WB assessments and the high number of labor days for a couple of countries (Azerbaijan with 275 and Indonesia with 195) may explain this finding.

Thirteen of the thirty PEFA assessments were co-financed in some capacity. Of the thirteen, three were financed from multi-donor trust funds already established in country. As a lead agency, the WB partnered with other organizations in six of the ten assessments, while the EC had for only four of the eleven assessments they led. However, when not in the lead, the EC contributed to four assessments and the WB only one. In this sample, DFID was the only PEFA partner to not lead an assessment but contribute financially.

There was an average of 3 assessors on the core PEFA team of EC led assessments those lead by other PEFA partner organizations. The WB averaged 5. The weighted average cost of an assessor, based on days per person, was \$1,068. The average daily rate of an EC assessor was \$1,285, and \$1,286 for assessors hired by other PEFA partner organizations that led an assessment. The average cost of a WB assessor was \$767. The lower rate for the WB may be due to several reasons including the less frequent use of

consulting firms, the more comprehensive accounting of labor inputs and the frequent tax exemption of salaries and fees of the assessors.

Size of country – in terms of population and to a slightly lesser extent total GDP - showed a significant correlation with the number of total labor days used to complete an assessment. It has not been possible to find any correlation between level of resource inputs and assessment report quality as measured by the Secretariat's compliance index.

The full details of the survey are presented in Annex B to this report.



## Chapter 5

### Conclusions and Recommendations

#### Roll-out

The number of PEFA assessment reports received by the Secretariat increased to 151 as of October 2009 from 67 at August 2007 reported in the previous monitoring report 2007. The number of countries covered increased to 102, the difference representing SNG and repeat assessments. The vast majority of assessments continue to be at central government level. The vast majority are stand-alone PEFA assessments, the proportion of stand-alone assessments having increased over the current reporting period, and the vast majority use the full PEFA Framework, again the proportion increasing during the current reporting period.

An increasing share of finalized reports is becoming available to the public through the internet (up from 42% in 2007 to 56% in October 2009).

Use of the Framework at SNG is still very limited, but is growing. Large countries, such as India, Pakistan, Argentina and Brazil have conducted SNG assessments. At present, the number of final reports is very small. In only two cases did assessors use the guidance on use of the PEFA Framework at sub-national government level.

#### Compliance

This report reviews the quality of 71 PEFA assessments received and reviewed by the Secretariat during the period under review, including 8 carried over from the previous period, representing initial drafts received but not revised/finalized until after April 1, 2007. Of the 71 reports, 44 are in revised/final draft stages. Sixty of the reports are at central government level. Twelve of the reports are repeats.

The average compliance index has steadily improved over time by around ten percent each fiscal year, starting with 56% in the first year of the application of the Framework (FY06) to 91% for FY09 (July-March only). The improvement is distributed across all indicators and core dimensions.

The difference in the compliance index between the initial draft and the revised/final draft is significant (59% to 80%), indicating that assessment teams have been judicious in addressing the comments of the PEFA Secretariat and others. The CI for the initial drafts of the 46 reports covered by the previous reporting periods (excluding the 8 carried over to the period under review) was 48 percent, indicating that the improvement in the final draft from the initial draft was somewhat higher than during the previous reporting period.

The number of “problem” indicators (defined as those with a compliance index below 60%) has markedly fallen, to only one indicator (PI-7) from 6 indicators identified in the previous report.

The Framework is being followed more closely, with the full indicator set usually being used. The incidence of “Not Used” and “Not rated” indicators has fallen markedly. Nevertheless, it is still the case that some indicators, particularly the Donor Practice indicators (mainly D-2-D-3) are not used when they could be used. . Not Used indicators mainly relate to the donor indicators, revenue administration indicators (e.g. Tuvalu) and PI-19 (e.g. Gambia, Pakistan) due to this being covered under an OECD-BIS assessment.

The diligence with which other sections of the Performance Report are completed also appears to have improved, most notably with regard to the Summary Assessment., though there is room for further improvement.

The improvement in the quality of the Performance Reports may be attributable to:

- Higher quality of the assessment teams, mainly the result of training received in the use of the Framework, the diminishing of the practice of the team consisting of only one person (typically for small island countries), and the experience gained through conducting a number of assessment reports (some consultants now have several PEFA assessment reports under their belt).
- Allocation of more time to conduct the assessments;
- The much greater involvement of governments in the assessment process, increasingly, albeit from a low base, in the form of self-assessments.
- Greater attention paid to quality assurance.

Nevertheless, there is still room for improvement. Twenty six initial drafts have yet to be revised/finalized (at least to the Secretariat’s knowledge). This might be expected for the six reports received between January and March 2009, but this still leaves twenty reports reviewed by the Secretariat prior to the end of last year, that have not been revised/finalized.

The number of revised drafts sent to the Secretariat for comment (including the answering by the Secretariat of queries from the assessment team or sponsoring donor agency) numbered only 7 out of the 43 reports. For quality assurance to work properly, the Secretariat should assess all revised reports before they become final drafts.

The quality of summary assessments and other sections of the assessment report appear to have improved since the last reporting period, though scope for improvement remains. Encouragingly, Secretariat comments on the initial drafts of these sections are addressed in the majority of reports. Preparation of summary assessments would benefit from strengthened analysis of the relative importance of the indicators where performance has been identified as being weak. This would help to point the way forward in terms of revising PFM reform strategies/action plans.

Practices in conducting PEFA assessments appear to have improved in all areas, with only a very small number of assessment reports impacted negatively by poor practices in terms of quality. Governments have become far more involved in assessments, including those assessments which are donor-led (still the vast majority). Practices that could be significantly improved include the provision of PEFA training (using the 2 or 3 day course prepared by the Secretariat), the shortening of the time taken to finalise the report, still too long in many cases, particularly in the cases where the PEFA assessment is part of a larger product (e.g. PEMFARs), and the lengthening of the time between assessments; the time should be a minimum of three years.

Provision of training increased sharply during the period under review, either by the Secretariat directly, or by other parties (e.g. the World Bank, and trainers “accredited” by the Secretariat). Training has been provided to governments, consultants and donor agencies. However, the number of training events for government officials, arguably the most important target for training, is still small though increasing.

Twelve repeat assessments were conducted during the reporting period. Out of these, only seven can be considered reasonably satisfactory in that they checked for comparability of scores and indicated what progress had been made since the previous report. Only one (Madagascar) can be considered fully satisfactory. A major issue is that the repeat assessment teams have probably not had access to the Secretariat’s comments on the first report. So, although comparability issues were identified, assessment teams may not have taken into account all the compliance issues raised in the first report.

At the initiation of the analysis in this monitoring report, only 3 years have lapsed since the first assessments were carried out. It is not surprising, therefore, that most assessments to date have not followed the recommended 3-5 years interval between baseline and repeat assessments. However, short intervals combined with frequent shifts in the leading agency could indicate that assessments are not done for commonly agreed purposes in a well-coordinated manner.

A survey of the costs/resource use in implementing a PEFA assessment was included as part of this monitoring report. It shows that the average costs and labor input are relatively uniform across regions and leading/funding agencies, though variance as regards the latter does exist partly as a result of different approaches to mobilizing the teams of assessors.

## **Recommendations**

- Proactive government involvement in the assessment process should continue to be promoted, including enhanced training effort for government officials to play a key role in assessment implementation.
- The standing recommendation of undertaking formal and full repeat assessments every 3-5 years, and certainly not annually, remains valid.

- Partners will encourage task teams to share concept notes at the draft stage with the Secretariat for comment
- The peer review process should be identified at the CN stage, quality assurance arrangement should be transparent and set out in the CN/TOR and explained in the full report
- For reports on which the PEFA Secretariat has provided comments, Partners are encouraged to share with the Secretariat the revised versions so as the Secretariat may follow-up on how the comments were addressed
- The Secretariat will develop guidelines for repeat assessments. The guidelines would include the recommendation to specifically set out in the CN/TOR the need for the assessment to track performance change since a specific earlier assessment.
- Repeat assessment teams should be provided Secretariat comments on the earlier finalized assessment report in order improve the basis for tracking progress.
- To strengthen Summary Assessments - focus more on the relative importance of weak linkages – guidance and training on formulating this section should be enhanced.
- PEFA training should be strengthened to include a module on the structure of the public sector and its importance for distinguishing national and sub-national level in assessments.
- The feasibility of developing a standard compliance index for CN/TORs and monitoring of its development as is currently the case for the compliance index for indicator assessment should be investigated.
- As a supplement to the compliance index for indicator assessment, a standard method of monitoring compliance of other parts of the PFM-PR should be development.
- PEFA assessment reports should disclose, as a standard, a statement on resource use in implementing the assessment and names of the assessment team members.

## Annex A

### PEFA Assessments used for assessing quality

<b>Applications</b>	<b>Date</b>	<b>Region</b>	<b>Lead agency</b>	<b>Type of assessment</b>	<b>Comments regarding PEFA Secretariat peer review</b>
Benin	Sep-07	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Botswana	Oct-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Burundi	Dec-08	AFR	EC	PFM-PR	Draft version: 1 Secretariat review of: draft
Cameroon	Nov-07	AFR	EC	PFM-PR	Draft version: 1 Secretariat review of: draft
Cape Verde	Nov-08	AFR	EC	PFM-PR	Draft version: 1 Secretariat review of: draft
Central African Republic	Jun-08	AFR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Comoros	Jan-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Cote d'Ivoire	Nov-08	AFR	WB	Integrated	Draft versions: 1 Secretariat review of: draft and final
Ethiopia	Oct-07	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Ethiopia-Regional govt.	Oct-07	AFR	EC	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft and final
Gambia	Mar-09	AFR	WB	Integrated	Draft version: 1 Secretariat review of: draft
Guinea	Jul-07	AFR	France	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Kenya	Oct-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Liberia	Dec-08	AFR	WB	Integrated	Draft versions: 2 Secretariat review of: both drafts
Madagascar	May-08	AFR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Malawi	Jun-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final

Mauritania	Mar-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Mauritius	Jun-07	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Mozambique	Feb-08	AFR	Norway	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Namibia	Jun-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Niger	Dec-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Rwanda	Jun-08	AFR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Sao Tome and Principe	Jun-07	AFR	WB	Integrated	Draft versions: 1 Secretariat review of: draft and final
Senegal	Dec-07	AFR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Senegal (Ville de Dakar)	Jan-09	AFR	France	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Seychelles	Jun-08	AFR	EC	PFM-PR	Draft versions: 2 Secretariat review of: both drafts
Sierra Leone	Dec-07	AFR	DFID	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
South Africa	Sep-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Tanzania	Jun-08	AFR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft
Togo	Nov-08	AFR	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Uganda	Mar-08	AFR	Self	PFM-PR	Draft versions: 0 Secretariat review of: final
Uganda	Dec-08	AFR	WB	PFM-PR	Draft versions: 2 Secretariat review of: both drafts
Zambia	Jun-08	AFR	Self	PFM-PR	Draft versions: 1 Secretariat review of: draft
Indonesia	Oct-07	EAP	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Papua New Guinea	Oct-08	EAP	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft

Philippines	Oct-07	EAP	WB	PFM-PR	Draft versions: 2 Secretariat review of: both drafts
Solomon Islands	Nov-08	EAP	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Thailand	Dec-08	EAP	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft
Tonga	Sep-07	EAP	AusAid	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Tuvalu	Apr-07	EAP	AsDB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Armenia	Oct-08	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Azerbaijan	Jan-08	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Georgia	Feb-08	ECA	WB	Integrated	Draft versions: 1 Secretariat review of: draft
Kosovo	Mar-07	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Moldova	Jun-08	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Montenegro	Feb-09	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Tajikistan	Jun-07	ECA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Belize	Oct-08	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Brazil-Ceara State	Feb-09	LAC	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Brazil-Distrito Federal	Feb-09	LAC	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Brazil-Minas Gerais State	Feb-09	LAC	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Brazil-Pernambuco State	Feb-09	LAC	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Dominica	Apr-07	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Dominican Republic	May-07	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final

Guyana	Dec-07	LAC	IADB	Integrated	Draft versions: 1 Secretariat review of: draft and final
Haiti	Jan-08	LAC	WB	Integrated	Draft versions: 1 Secretariat review of: draft and final
Jamaica	May-07	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Paraguay	Apr-08	LAC	WB	Integrated	Draft versions: 1 Secretariat review of: draft
Peru	Dec-08	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Trinidad and Tobago	Oct-08	LAC	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft
Iraq	Mar-08	MENA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft
Jordan	Apr-07	MENA	EC	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Morocco	Feb-09	MENA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft
Yemen	Jun-08	MENA	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Norway	Jun-08	Other	Norad	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
Switzerland-Lucerne Canton	Jan-09	Other	SECO	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft
Afghanistan	Jun-08	SAR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft and final
India-Himachal Pradesh State	Mar-08	SAR	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft and final
India-Maharashtra State	Mar-08	SAR	WB	PFM-PR-SN	Draft versions: 1 Secretariat review of: draft and final
Pakistan-Sindh Province	Mar-08	SAR	WB	PFM-PR-SN	Draft versions: 2 Secretariat review of: 1st draft
Pakistan	Mar-09	SAR	WB	PFM-PR	Draft versions: 1 Secretariat review of: draft



## Annex B

### Survey of the Costs of Implementing PEFA Assessments

#### B.1 Introduction

The PEFA framework was designed to be used as a monitoring tool to measure country PFM performance over time. The process for undertaking an assessment and writing of the PFM-Performance Report is usually quite standardized even if variations do occur, especially in connection with integrated analytical products. The question has often been asked: What does it cost to implement a PEFA assessment? This survey attempts to give some answers to this question and to explore variations between costs as a result of who leads the assessment, what kind of country the assessment is for and whether the resource input affects the assessment quality.

#### B.2 Context and Approach

The data collection was centered on the development of a standard set of questions and filling them out based on information collected from Concept Notes, Terms of References (TOR), PEFA assessments, task team leaders, operations analysts, and others with access to project expenditure information. The standard set of questions was as follows:

- 1.) What was the total budget envelope or expenditure?
- 2.) Did the lead agency finance the entire assessment? If not, what was the contribution of other organizations involved?
- 3.) What was the composition of the team? How many persons were on the assessment team, what was their position title, and how many years of experience was requested that they have?
- 4.) What was the breakdown of labor (days allocated per position and daily rate) and other costs (per diem, airfare, transportation, et cetera)?

This data was sought for the same 71 assessment that form the basis for assessing quality issues in this Monitoring Report. The data available was obtained in four distinct forms:

- Extracted from a concept note, terms of reference, or the assessment report,
- Submitted in an e-mail from the Assessment Task Leader with answers to the above questions,
- Extracted from a form signed by the consulting firm undertaking the assessment indicating the maximum labor and reimbursable costs (per diem, travel, translation, et cetera) that could be charged (in the case of the EC),
- Extracted from disaggregated financial data from an online database/accounting system (in the case of the WB).

The decision to include an assessment in this analysis was based on the availability of the budget envelope or actual cost and its composition. That means that an assessment was not included if:

- Only the labor or other costs were included but not both
- PEFA budgeted or actual costs were mixed with those of a larger project or analytical work
- Incomplete PEFA budgeted or actual cost information received due to contributions made from multiple donors and IFIs and the uncertainty of how much was financially contributed on their behalf

The World Bank for example often conducts PEFA assessments under the work program of a larger project without creating a separate work order or billing code; as a result, it was difficult to determine the amount of labor used and other related expenditures for the PEFA assessment work. For those assessments with contributions from numerous donors, the lead agency was often able to provide a budget but the other contributor(s) was only able to provide a description of the contribution, such as covering the labor and other expenses of a local consultant or sponsoring training.

Of the 71 assessments considered for the monitoring report and this cost study, the requisite data could be obtained for 30 assessments, which were used for most of the proceeding analyses. This includes 15 led by the EC, 10 by the WB, and 5 by other PEFA partner organizations (France, Norad and SECO) - see table B-10 in the annex for the complete list of assessments.

When information was not available a N.S. (not specified) is presented in the tables below or recognizable by the fewer number of assessments analyzed. For example, the total labor days used for four countries were not available so they were excluded (see table 2 as a reference). The average daily labor rate analysis is the only one that includes some gap filling. An explanation of how the gap filling was done is presented in that section.

### **B.3 The Overall Cost of a PEFA Assessment**

The analysis conducted in this study on the overall cost of a PEFA assessment comes with some considerations. First, the labor hours spent for peer review were not included because the majority of the reviewers did not code their time separate from other daily/project tasks. Second, the totals were converted to U.S. Dollars when in another currency, using the exchange rate of the first day of the month the main mission started.<sup>7</sup> Third, countries are included that have some financial contributions missing but were

---

<sup>7</sup> Main mission is defined as the first mission in which most of the data collection, analysis, and writing is done in country. It is different than the scoping mission, which would come before and may include a workshop and decision trip to determine if an assessment will be pursued, or follow-up mission, which would come after if further data collection efforts were needed or a closing workshop.

considered insignificant (see table B-5 under the section on co-financing). Fourth, due to the limited sample size, the financial figures used are a mixture of budget and actuals. For most of the EC sponsored assessments the ceiling of a firm implementing the assessment could charge was used, while for the WB and other lead organizations, the actuals were used. Table B-1 below provides the mean, median, and range of the overall cost of PEFA assessments by region and global.

*Table B-1: Average Cost of a PEFA Assessment by Region and Global*

<b>Region</b>	<b>No. of assessments</b>	<b>Mean</b>	<b>Median</b>	<b>Range</b>	
AFR	17	\$ 129,624	\$ 118,279	\$ 25,206	\$ 272,281
EAP	2	\$ 93,801	\$ 93,801	\$ 67,282	\$ 120,320
ECA	3	\$ 123,098	\$ 91,499	\$ 85,675	\$ 192,119
LAC	3	\$ 147,062	\$ 83,124	\$ 77,256	\$ 280,805
MENA	2	\$ 185,920	\$ 185,920	\$ 106,980	\$ 264,860
SAR	1	\$ 112,000	\$ 112,000	\$ 112,000	\$ 112,000
Other	2	\$ 51,287	\$ 51,287	\$ 41,250	\$ 61,324
Global	30	\$ 126,270	\$ 109,490	\$ 25,206	\$ 280,805

Since over 50 percent of the assessments used in this data set are in Africa, it is not surprising that the global mean, median and range are close to those of Africa. Nevertheless, the rest of the regions (apart from “Other”) are not too far off from the global averages. The most significant differences are with Other. The Norwegian and subnational assessment for the Swiss canton Lucerne are the only assessments included in the Other category in table B-1. Zero non-labor related costs, and lower than average labor rates resulted in a below average price tag. In Africa, two assessments fall under USD 80,000, in Guinea and Benin.

*Table B-2: Average Days of Labor of a PEFA Assessment by Region and Global*

<b>Region</b>	<b>No. of assessments<sup>8</sup></b>	<b>Mean</b>	<b>Median</b>	<b>Range</b>	
AFR	15	82	78	30	146
EAP	2	115	115	35	195
ECA	3	142	83	64	275
LAC	2	89	89	48	130
MENA	2	100	100	62	138
Other	2	69	69	57	80
Global	26	92	78	30	275

<sup>8</sup> Total labor days were not available for 4 assessments, thus only twenty-six assessments are considered.

Table B-2 provides a similar picture as table B-1, the mean and median of the Africa region are close, if not the same, as the Global mean and median. Two noteworthy differences are with EAP, where the mean and median are below the Global averages in overall cost but greater than the Global averages in days of labor, and ECA, where the mean is similar to the Global mean in overall cost but significantly above in overall days of labor. In both cases, the high range figures, 195 days (EAP) and 275 days (ECA) explain this outcome.

Table B-3 provides the mean, median, and range of the overall cost of PEFA assessments by lead agency and table B-4 by total days in labor.

*Table B-3: Overall Cost of a PEFA Assessment by Lead Agency*

Lead Agency	No. of assessments	Mean	Median	Range	
EC	15	\$ 137,876	\$ 103,333	\$ 50,164	\$ 280,805
WB	10	\$ 134,492	\$ 114,575	\$ 85,675	\$ 264,860
Other	5	\$ 75,007	\$ 61,324	\$ 25,206	\$ 128,975

The EC and the WB, the two agencies that have completed the most assessments (over 85% as of March 2009), also have similar mean and median cost. However, the mean and the median of the assessments completed by other agencies are significantly lower. What might explain this difference? While two of the assessments completed, Senegal – Ville de Dakar (\$118,279) and Mozambique (\$128,975) are close to the Global averages; Guinea (\$25,206), Norway (\$41,250) and Switzerland-Lucerne Canton (\$61,324) are not. Guinea may be an anomaly, while the zero and low non-labor costs of Norway and Switzerland-Lucerne Canton explain their lower overall costs as noted above.

*Table B-4: Overall Days of Labor of a PEFA Assessment by Lead Agency*

Lead Agency	No. of assessments	Mean	Median	Range	
EC	14	79	72	30	146
WB	8	123	87	62	275
Other	4	76	72	57	104

In this sample, the WB has a significantly higher mean than the EC and Other, the range shows a greater spread as well. This may be related to the WB leading assessments in some of the largest countries, such as Indonesia and Pakistan, ref further analysis in section B-6. Relationships of this nature are discussed later on in the report in the section on correlations.

## **B.4 Co-financing arrangements**

For this study, co-financing refers to an assessment that received financial or in-kind support from more than one organization. In-kind support does not include peer reviews

but does include the financing of an assessor or training for persons involved in the assessment. Thirteen of the thirty assessments were co-financed in some capacity. Of the thirteen, three were financed from multi-donor trust funds already established in country. Table B-3 below provides a summary of the thirteen co-financed assessments with the contribution of each partner recognized. N.S. (not specified) is marked if a financial amount was not available but a contribution was made. The column “Other” includes donors and the PEFA Secretariat.

*Table B-5: Co-financed PEFA assessments (in USD)*

Country	Lead Donor	Lead	AfDB	AsDB	DFID	EC	W B	Other	Budget Envelope/Actual
Benin	EC	46,019						4,145	50,164
Guinea	France	20,287				4,919			25,206
Kenya	EC	103,333						N.S.	103,333
Madagascar	WB	120,000	15,000			15,000			150,000
Mozambique	Norway	128,975						N.S.	128,975
Senegal*	WB	117,150							117,150
Senegal (Ville de Dakar)*	France	118,279							118,279
Uganda	WB	27,390			22,157			54,772	104,319
Indonesia*	WB	120,320							120,320
Solomon Islands	EC	67,282						N.S.	67,282
Peru	EC	280,805					N.S.		280,805
Morocco	WB	196,042				68,818			264,860
Pakistan	WB	40,000		6,000	60,000	6,000			112,000

\* Financed from a multidonor trust fund

N.S.: Contribution not specified in monetary amount.

It is also worth mentioning the tendencies of organizations to pursue a co-financing arrangement. As a lead agency, the World Bank has partnered with other organizations in six of the ten assessments, while the European Commission has for only four of the eleven assessments they led. However, when not in the lead, the European Commission has contributed to four assessments and the World Bank only one. In this sample, DFID was the only PEFA partner to not lead an assessment but contribute financially.

## **B.5 Composition of an Assessment Team**

Since a PEFA assessment covers the entire gamut of a government’s PFM system, forming a team with a diverse skill set and adequate level of experience facilitates the process in producing a quality assessment. While this analysis is not concerned with the skill set of each assessment team member, the number of years of experience requested in the TOR or concept note and the cost of a consultant are considered. In most assessments led by the World Bank, determining the composition of the core assessment team was difficult to ascertain since numerous persons are listed in a report as contributors.

Nevertheless, an attempt was made by triangulating names listed in concept notes, the report and in the project expenditure information. Table B-6 below lists the total number of core members of an assessment team. This excludes translators, reviewers, and others who contributed at the margin.

*Table B-6: Composition of a PEFA assessment team – Total number of core assessors*

<b>Lead Agency</b>	<b>No. of assessments</b>	<b>Mean</b>	<b>Median</b>	<b>Range</b>	
EC	15	3	2	2	4
WB	10	5	4	3	11
Other	5	3	3	2	4
Global	30	3	3	2	11

While the mean and the median for the EC, and Other are similar to the Global averages, the World Bank is slightly higher. An even greater difference exists in the range where up to eleven members were part of an assessment team (Pakistan). The geographic size and large population may be related to this situation.

Similarly, World Bank concept notes infrequently mention the number of years of experience an assessment member should have to be part of the team; therefore only two assessments are included in the years of experience analysis. On the other hand, the EC lists the number of years of experience in all of their TORs but does not mention the person(s) who will be part of the assessment team. This is due to the fundamental difference in how each organization forms an assessment team. The WB will often use a combination of in house staff and consultants that may or may not be chosen based on an open tender process<sup>9</sup> while the EC uses only consultants who are provided by the company that is chosen to undertake the assessment.

Table B-7 below shows the weighted average expressed as a percentage for each organization. For example, of the 15 assessments lead by the EC, in 44% of the assessments it indicated in the TOR that an assessor should have 15 years of experience or more.

*Table B-7: Composition of a PEFA assessment team – Years of experience requested of core assessors*

<b>Lead Agency</b>	<b>No. of assessments</b>	<b>≥15</b>	<b>≥10</b>	<b>≥5</b>
EC	15	44%	42%	14%
WB	2	29%	46%	25%
Other	4	69%	15%	17%
Global	21	47%	37%	16%

<sup>9</sup> While all consultants are originally selected based on an open tender, a contract up to 150 days per fiscal year can be renewed an unlimited number of times as long as there is not a significant deviation from the Terms of Reference.

The weighted average cost of an assessor is expressed in Table B-8 below, organized by lead agency. The cost of each assessment includes the weighted average, based on days, of the core assessors. It combines international and local consultants. In four cases (Azerbaijan, Peru, South Africa, Uganda), the number of labor days were indicated in a concept note, TOR, or project expenditure document, but the rate was not. To compensate, the average of each organization was used and are included in the overall averages expressed below. In two of the cases (Armenia and Morocco), not all of the daily rates of assessors were provided but some were, therefore, the weighted average rate of the consultants whose rates were provided at the same level was used.

*Table B-8: Composition of a PEFA assessment team – Average daily labor rate*

<b>Lead Agency</b>	<b>No. of assessments</b>	<b>Mean</b>	<b>Median</b>	<b>Range</b>	
EC	14	\$ 1,285	\$ 1,285	\$ 1,041	\$ 1,638
WB	8	\$ 767	\$ 748	\$ 372	\$ 1,337
Other	4	\$ 1,286	\$ 885	\$ 516	\$ 1,361
Global	26	\$ 1,068	\$ 1,165	\$ 372	\$ 1,638

The mean and median daily labor rate for EC consultants is over 50% higher than those of the WB. There may be three reasons for this difference. Firstly, the WB uses primarily own staff and individual consultants (who are largely treated as staff) rather than consultants from companies. Some company overheads and profit will therefore be included in EC rates but not in WB rates. Secondly, many WB staff/consultants receive salaries and consulting fees net of tax. Consulting fee rates paid by the EC through companies will always be taxable. The difference between fee rates gross and net of tax for the WB corresponds largely to the difference in average labor day costs. Furthermore, in this sample the WB used local consultants for five assessments and the EC only one. The latter aspect brought down the average rates considerably in these 5 cases, or 63% of the WB sample. The highest local daily labor rate of a WB assessor was 400 USD.

## **B.6 Explanatory factors**

Is there a relationship between the cost of an assessment and country characteristics such as size, measured by GDP or population? Table B-9 below provides the correlation coefficients ( $r$ ) of the data set in this sample and the GDP and population of the assessed countries.

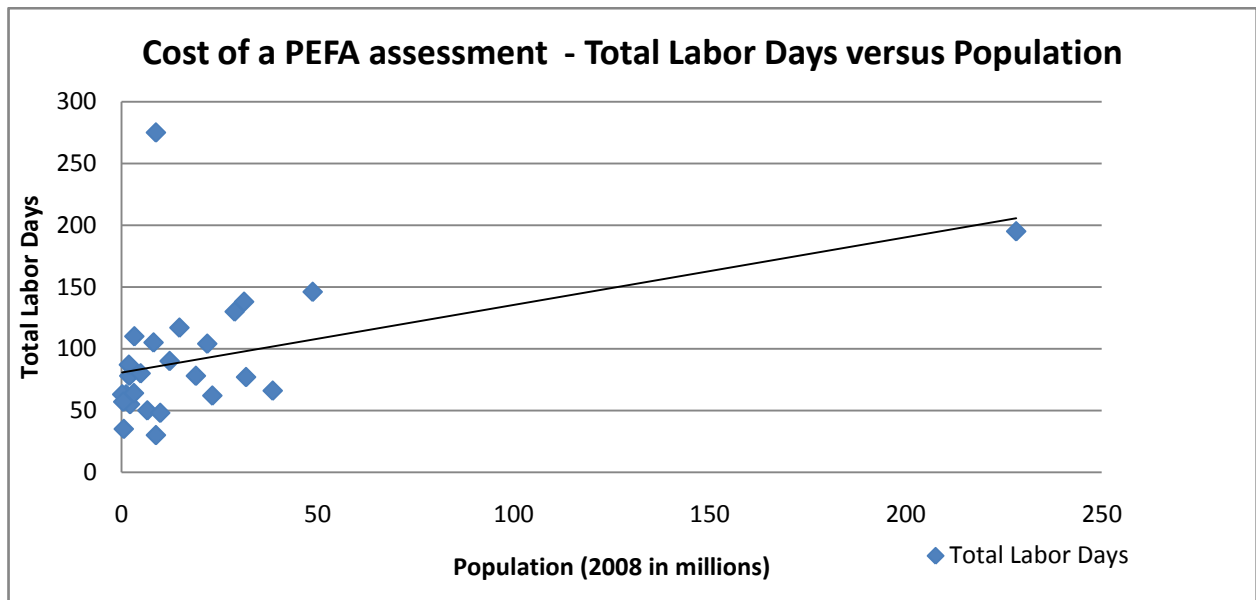
Table B-9: Correlation Coefficients of cost data, GDP and population

	Population (2008 in millions)	GDP (2008 in billions)	Compliance Index (initial)
Budget Envelope/Actual	<b>0.11</b>	<b>0.11</b>	<b>0.13</b>
Total Labor Days	<b>0.46</b>	<b>0.41</b>	<b>-0.10</b>

The correlations of interest are highlighted in bold in table B-9. In this sample, there is no relationship between the budget envelope or actual cost and population ( $r = 0.11$ ), GDP (0.11) and the compliance index of draft reports (0.13). In addition, there is no relationship between total labor days and the compliance index of draft reports (-0.10). However, there is a moderately positive relationship between total labor days and population (0.46) and GDP (0.41). After testing the significance of  $r$  at the 0.05 level of significance for population (-7.07) and GDP (-1.04) the relationship is considered significant for population but spurious for GDP. Four countries were removed from the part of the analysis using total labor days, Belize, Guinea, Madagascar, and Pakistan since the total labor days were unavailable.

Graph B-1 below plots total labor days and population. The two outliers are Azerbaijan in the upper left hand corner with a population of 8.7 million and 275 total labor days and Indonesia with a population of 228.2 and 195 labor days. However, the graph does not change significantly if those two outliers are excluded from the plot.

Graph B-1: The relationship between total labor days and population of a country





## B.7 Conclusion

While the sample data for this cost study may contain a few caveats, such as mixing budgeted with actual cost data, and the reduced sample sizes for some analysis, such as the years of experience the lead agency would request assessment team members have, some robust findings have emerged as follows:

- The overall cost of a PEFA assessment is on average in the order of USD 126,000, but wide a wide range from USD 25,000 to USD 280,000.
- the number of labor days used is in average about 92 with a range of 30-275, but typically in the range of 75-100 days
- The assessment costs in USD are very similar for the World Bank and the EC, whereas the assessments led by bilateral donor agencies have been somewhat lower. However, measured in labor day inputs EC and bilateral agency assessments are very similar whereas the World Bank has used about 50% more labor days per assessment.
- The core assessment team typically consists of 3 persons, and have 10 to 15 years of experience, if not more.
- Average cost per labor day is USD 1100-1,300 per day for assessors funded by EC and bilateral agencies, but only USD 767 per day for the World Bank
- Size of country – in terms of population - showed a significant correlation with the number of total labor days used to complete an assessment.
- It has not been possible to find any correlation between level of resource inputs and assessment report quality as measured by the Secretariat’s compliance index.

These observations should be taken not as absolutes but reference points for assessment managers. A more robust study might include a separation between the budget envelope and actuals, more comprehensive data on co-financing, and further definition to who should be considered part of the “core” assessment team.

<b>Table B-10 Data from the 30 Countries used for the study</b>	<b>Region</b>	<b>Population (2008 in millions)</b>	<b>GDP (2008 in billions)</b>	<b>Lead Donor</b>	<b>Budget Envelope/Actual</b>	<b>Total Labor Days Used</b>	<b>Total Number in Core Team</b>	<b>Average Daily Labor Cost of Assessor</b>
<b>Country</b>								
Benin	AFR	8.7	6.7	EC	\$ 50,164	30	3	\$ 1,115
Botswana	AFR	1.9	13	EC	\$ 173,427	78	2	\$ 1,638
Burundi	AFR	8.1	1.2	EC	\$ 134,539	105	3	\$ 1,041
Cameroon	AFR	18.9	23.4	EC	\$ 148,293	78	2	\$ 1,432
Guinea	AFR	9.8	4.3	France	\$ 25,206	N.S.	3	N.S.
Kenya	AFR	38.5	34.5	EC	\$ 103,333	66	4	\$ 1,101
Madagascar	AFR	19.1	9	WB	\$ 150,000	N.S.	4	N.S.
Mauritania	AFR	3.2	2.9	EC	\$ 197,002	110	2	\$ 1,353
Mozambique	AFR	21.8	9.7	Norway	\$ 128,975	104	3	\$ 772
Namibia	AFR	2.1	8.6	EC	\$ 97,164	55	2	\$ 1,343
Niger	AFR	14.7	5.4	EC	\$ 195,012	117	3	\$ 1,327
Senegal	AFR	12.2	13.2	WB	\$ 117,150	90	3	\$ 717
Senegal (Ville de Dakar)	AFR	1.1	5.0	France	\$ 118,279	63	2	\$ 1,361
Seychelles	AFR	0.1	0.8	EC	\$ 100,987	63	3	\$ 1,235
South Africa	AFR	48.7	276.8	EC	\$ 272,281	146	3	\$ 1,285
Togo	AFR	6.5	2.8	EC	\$ 87,479	50	2	\$ 1,215
Uganda	AFR	31.7	14.5	WB	\$ 104,319	77	3	\$ 731
Indonesia	EAP	228.2	514.4	WB	\$ 120,320	195	8	\$ 372
Solomon Islands	EAP	0.5	0.6	EC	\$ 67,282	35	2	\$ 1,248
Armenia	ECA	3.1	11.9	WB	\$ 85,675	64	4	\$ 819
Azerbaijan	ECA	8.7	46.3	WB	\$ 192,119	275	4	\$ 625
Kosovo	ECA	1.8	3.2	WB	\$ 91,499	87	3	\$ 772
Belize	LAC	0.3	1.4	EC	\$ 77,256	N.S.	2	N.S.
Dominican Republic	LAC	9.8	45.8	EC	\$ 83,124	48	2	\$ 1,373
Peru	LAC	28.8	127.4	EC	\$ 280,805	130	4	\$ 1,285
Morocco	MENA	31.2	86.3	WB	\$ 264,860	138	7	\$ 1,337
Yemen	MENA	23.1	26.6	WB	\$ 106,980	62	3	\$ 764
Pakistan	SAR	166	168.3	WB	\$ 112,000	N.S.	11	N.S.
Norway	Other	4.8	450	Norad	\$ 41,250	80	2	\$ 516
Switzerland-Lucerne Canton	Other	0.36	15.8	SECO	\$ 61,324	57	4	\$ 998