

Tracking Progress in the Quality of PFM Systems in HIPC

An update on past assessments using PEFA data

Paolo de Renzio and Bill Dorotinsky*

November 2007

* Paolo de Renzio is a doctoral student at the University of Oxford and Bill Dorotinsky is a Deputy Division Chief in the International Monetary Fund Fiscal Affairs Department. The views expressed in this paper are solely the authors, and do not necessarily reflect the views of the IMF or the PEFA program.

The background research for this paper was financed by the PEFA Program (PEFA is a partnership program of the World Bank, the European Commission, the UK's Department for International Development, the Swiss State Secretariat for Economic Affairs, the French Ministry of Foreign Affairs, the Royal Norwegian Ministry of Foreign Affairs, the International Monetary Fund and the Strategic Partnership with Africa) and benefited from the hospitality and support of the PEFA Secretariat in the period from July to October 2007.

Executive Summary

In recent years, developing countries and donor agencies have recognized the importance of public financial management (PFM) systems for achieving national objectives, including economic growth and poverty reduction. The joint IMF-World Bank work on tracking poverty-reducing public spending in Heavily Indebted Poor Countries (HIPCs) has shown the feasibility of using indicator-based assessment instruments to more consistently evaluate and track country PFM system performance over time. The Public Expenditure and Financial Accountability (PEFA) methodology, launched in 2005 and now being rolled out worldwide, further developed this approach with a more comprehensive set of indicators and a refined assessment scale.

This paper assesses changes in PFM system performance in HIPCs from 2001 to 2006 using data from the original HIPC expenditure tracking methodology, updated with additional information from PEFA assessments conducted after the last round of HIPC assessments in 2004. It covers 11 of the 16 HIPC indicators and 15 of the 26 countries from the original assessment in 2001. Results are presented according to the number of HIPC benchmarks met and to the raw numerical scores for each indicator. In addition to presenting the results along the original dimensions of phases of the budget cycle (formulation, execution, reporting), the paper also presents the results using a new categorization (quality of budget information, the budget as policy instrument, and effectiveness of control and oversight functions).

The results must be interpreted with caution. As noted above, only 11 of the original 16 indicators could be tracked. Three of the indicators that could not be tracked are in budget reporting, reducing the coverage of the indicator sample to only part of the over-all budget cycle, and reducing the likelihood of detecting significant changes in PFM system performance. Only 15 of the original 26 countries could be tracked over the period, a significantly smaller sample size. And in retrofitting 2005-7 PEFA assessment results to the HIPC methodology, some degree of judgment was necessary. The limited sample size, reduced indicator set, and methodological issues all reduce the ability to draw strong conclusions or generalize the results. Countries may have made progress or introduced reforms in PFM areas not covered by the indicator sets.

The actual results suggest limited and uneven progress across the countries covered. Results for 2004-6 show a slightly more marked improvement in PFM systems than 2001-2004 across the 11 indicators. More specifically:

- For the 11 indicators covered, 2001-2004 results for the 15 countries show *no change* or a *slight improvement* in PFM system performance. Four countries improved the number of benchmarks met, four declined, and seven remained constant. The total number of benchmarks met for the pool of countries remained constant over 2001-2004. Raw results (changes in raw score, irrespective of benchmarks met) for the

- same period show nine countries improving, four declining, and two remaining unchanged.
- For the period 2004-6, using the original approach, the results show a further *slight improvement* compared to 2001-2004. In terms of benchmarks met, seven countries improved, five declined, and three remained constant. The total number of benchmarks met for the entire pool of 15 countries improved slightly. In terms of raw scores, eight countries improved, five declined, and two remained constant.
 - In terms of trajectories of change, most countries showing improvements in PFM system performance tended to continue to improve, and about half of those on a declining trend reversed course (to reach 2001 levels or better), at least over the limited time horizon under review.¹

Looking at the whole period from 2001 to 2006, again the overall picture hides significant variation across countries. Looking at benchmarks met, five countries showed an improvement over the 2001-2006 period (with increases of up to six benchmarks for Ghana), six experienced a decline and four remained largely unchanged. Looking at raw scores, the picture improves, with eight countries improving, four declining and three presenting no change in performance. The group of better performers remains similar but expands over time, with Burkina Faso, Tanzania, Ghana and Guyana leading both in terms of benchmarks met and raw scores. Poor performers include Benin, Guinea, Malawi and São Tomé & Príncipe. Changes across areas and indicators also reveal interesting patterns. Improvements are evident, for example, in budget classification, budget reliability and external audit. Areas where weaknesses persist include internal controls, the inclusion of donor funds in the budget and expenditure arrears.

The results highlight the need to better understand the dynamics leading to changes in PFM system performance. An initial attempt to explore dynamics suggests that initial PFM system condition and government commitment (measured through the implementation of action plans) are the best predictors of future improvements in assessed performance.

In the future, similar updates should be done on the basis of PEFA assessments and methodology, as PEFA reports constitute a much improved source of data and information, covering additional aspects of PFM (such as, for example, taxation and transparency) for a wider set of countries than just HIPC's.

¹ All four countries that improved in number of benchmarks met over 2001-2004 continued to show improvement over 2004-2006. Over half of countries showing no change over 2001-2004 showed declining performance over 2004-2006. On the other hand, half of the countries declining over 2001-2004 reversed course and showed improvement in 2004-2006. The trends were similar for raw scores, with two-thirds of countries showing improved raw scores over 2001-2004 continuing to improve over 2004-2006, both countries that showed no change over 2001-2004 declining over 2004-2006, and again half of the countries declining over 2001-2004 reversing course over 2004-2006.

1. Introduction

The quality of public financial management (PFM) systems is a key determinant of government effectiveness. The capacity to direct, manage and track public spending allows governments to pursue their national objectives and account for the use of public resources and donor funds. In recent years, developing countries and donor agencies have recognized the central importance of PFM systems for achieving national objectives, including economic growth and poverty reduction. Donor agencies have been devoting increasing attention and resources to assessing and strengthening the quality of PFM systems in recipient countries, both as a way to address fiduciary concerns, and to identify strengths and weaknesses that can affect achievement of national objectives, development performance and aid effectiveness.

Improvements in the ability to track PFM system performance over time is an important advance in the field of public financial management. The joint IMF-World Bank HIPC expenditure tracking work has shown the feasibility of using indicator-based assessment instruments to more consistently evaluate and track country PFM system performance over time. The PEFA instrument further developed this approach with a more comprehensive set of indicators and a refined assessment scale. For countries, these tools allow for the internal monitoring of their own PFM system performance, enabling better risk management and more targeted PFM reform strategy development. For external parties, they enable better targeted support to PFM reforms, and improved learning of what reforms work, under what conditions.

Occasional stock-taking of PFM reform progress across countries is a useful way of evaluating PFM reform progress more generally, and learn lessons about what is working. The series of Fund-Bank Board papers between 2001 and 2005 provided valuable inputs into learning about reform dynamics and directing technical assistance.²

This paper assesses the status of PFM systems in selected HIPC countries, and tracks progress in their performance from 2001 through to 2006. It is an update on the 2001 and 2004 HIPC assessments using information generated through more recent PEFA assessments. It covers the sub-set of countries and of indicators for which a complete dataset is available. These results should be of interest to the Executive Boards of the Bank and the Fund, to operational staff in both institutions, to the PEFA partners and other donor agencies involved in supporting PFM reforms, and to government officials, parliamentarians, civil society and the media in the countries included in the analysis.

² Papers included: *Tracking of Poverty-Reducing Public Spending in Heavily Indebted Poor Countries (HIPCs)* (March 2001); *Actions to Strengthen the Tracking of Poverty-Reducing Public Spending in Heavily Indebted Poor Countries (HIPCs)* (March, 2002); *Update on Implementation of Action Plans to Strengthen Capacity of HIPCs to Track Poverty-Reducing Spending* (March 2003); and, *Update on the Assessments and Implementation of Action Plans to Strengthen Capacity of HIPCs to Track Poverty-Reducing Public Spending* (April 2005).

The paper is organized as follows: a Background section provides some historical information summarizing the results of previous overviews; a Methodology section describes issues related to mapping the HIPC expenditure tracking indicators to PEFA indicators, the sample of countries, and related issues; a Results section summarizes the results of the analysis, tracking performance progress across countries, dimensions and indicators; and a Conclusions section summarizes key points and presents suggestions for further research.

2. Background and Summary of Previous Findings

In 2000, in the context of the Highly Indebted Poor Countries (HIPC) debt relief initiative, the Executive Boards of the World Bank and the International Monetary Fund asked staff to launch a series of PFM assessments. The assessments were meant to gauge the capacity of countries benefiting from debt relief to track poverty-reducing public spending, in order to ensure that additional resources were spent effectively to reach poverty reduction objectives. The main purposes of the assessments were to:

- assess the existing capacity for tracking overall public spending, including poverty-reducing spending, in the context of debt relief; and
- clarify what donor and technical assistance should be provided to improve systems for managing poverty-reducing and other public spending.³

These assessments were carried out in 2001 and 2004, and the results have been summarized in a Board Paper dated April 2005⁴, covering 26 countries and 16 indicators and benchmarks looking at different aspects of the budget process (Budget Formulation, Budget Execution, Budget Reporting and Procurement). The assessments also included the identification and follow-up of specific actions to address PFM weaknesses, and a summary of donor activities in support of PFM reforms.⁵

The 2005 IMF-World Bank Board paper reported that the total number of benchmarks met by the entire sample of 23 countries assessed improved over the period 2001-2004. The total number of benchmarks met increased from 137 in 2001 to 150 in 2004 (a 10 percent increase), while the average number of benchmarks met increased from 6 to 6.5. Four countries met nine or more benchmarks in 2004, compared to only one in 2001.

³ *Country Assessment and Action Plan (AAP) for HIPCs*. Revised Guidelines, October 2003 (<http://www1.worldbank.org/publicsector/pe/FinalHIPCAAPGuidance2003-04.pdf>)

⁴ *Update on the Assessments and Implementation of Action Plans to Strengthen Capacity of HIPCs to Track Poverty-Reducing Public Spending* (April 2005). (<http://www1.worldbank.org/publicsector/pe/HIPC/HIPCBoardPaperApril2005.pdf>)

⁵ All papers are available at <http://go.worldbank.org/6NCYI7K2V0>.

While overall country PFM system performance generally improved, there was variability in performance among countries. Five of the 23 countries reassessed in 2004 showed improvements in their PFM systems, meeting an additional three or more benchmarks compared to 2001. Ghana improved most, meeting seven benchmarks in 2004, up from one in 2001. Cameroon, Mali, Ghana, Senegal and Tanzania improved by three benchmarks each. The number of benchmarks met remained unchanged for Benin, Guinea, Rwanda, São Tomé and Príncipe and Zambia, and deteriorated by one benchmark or more for Bolivia, Gambia, Honduras, Madagascar, Mozambique, Malawi, Chad and Uganda.

Within countries, there was performance variance among different PFM areas and indicators. Of the three PFM areas assessed in both 2001 and 2004, budget reporting showed the largest improvements in benchmarks met, with improvements seen in the quality of fiscal reports, routine booking of transactions after the end of the fiscal year, and the timeliness of audited financial information. Budget formulation improvements included budget comprehensiveness, classification, identification of poverty-reducing spending, and integration of medium-term expenditure projections. Budget execution improvements occurred in the routine reconciliation of fiscal and banking records. Areas of weaker performance included comprehensiveness of fiscal operations coverage, multi-year planning, ability to manage expenditures (budget outturn and arrears), internal controls, timely expenditure reporting, and inclusion of aid in the budget and in budget reports.

After the second assessment in 2004, the HIPC expenditure tracking methodology served as an important input into a multi-donor effort to develop a common framework to assess country PFM systems. A number of bilateral and multilateral agencies agreed to adopt a unified and strengthened approach to supporting PFM reforms in low income countries.⁶ The Public Expenditure and Financial Accountability (PEFA) initiative⁷ is a partnership of eight entities⁸ which developed a PFM Performance Measurement Framework which has been used in more than 50 countries since its launch in June 2005. It is based on a set of 28 high-level indicators measuring PFM systems performance along six critical dimensions: (a) credibility of the budget; (b) comprehensiveness and transparency; (c) policy-based budgeting; (d) predictability and control in budget execution; (e) accounting, recording and reporting; and (f) external scrutiny and audit. Three additional indicators on donor performance are also included.

Important differences between the HIPC and PEFA approaches need to be recognized. HIPC expenditure tracking assessment followed a systems-approach, looking at

⁶ The Strengthened Approach is based on lessons learned in successful PFM reforms in countries across the globe, and has three pillars: (1) country-led PFM reform program; (2) country-level donor coordination around the government reform program; and (3) a common framework for assessing PFM reform progress over time.

⁷ www.pefa.org

⁸ The World Bank, the European Commission, the UK's Department for International Development, the Swiss State Secretariat for Economic Affairs, the French Ministry of Foreign Affairs, the Royal Norwegian Ministry of Foreign Affairs, the International Monetary Fund and the Strategic Partnership with Africa. There is also strong collaboration with the OECD/DAC Joint Venture on PFM.

critical elements of the overall PFM system. The HIPC approach focused on the capacity of low-income country PFM systems to track and report on poverty-reducing spending, including both a country's own resources and resources freed-up through debt relief. For this purpose, for each indicator, a 'benchmark' performance level was identified. Taken in total, if a country were able to meet the benchmark for all indicators, it would be deemed able to reliably track and report on poverty-reducing spending. The PEFA approach also follows a systems approach, but includes a broader set of indicators covering more aspects of the PFM system. The PEFA framework is meant to provide an overall assessment of the quality of PFM systems, without defining any specific benchmarks. While the HIPC assessment was targeted towards low-income country PFM systems, the PEFA framework is developed to be broadly applicable to all countries, regardless of their level of development, embedding international good practice standards.

3. Methodology

There are some important differences between the HIPC assessment methodology and the PEFA framework which affect the content of this paper as compared to the previous Board Papers summarizing and comparing the results of the 2001 and 2004 HIPC assessments. These relate mostly to the number of countries and indicators covered, and to the scoring and aggregation methods used.

Countries and Indicators

There are 15 of the original HIPC AAP countries for which an update is currently possible. The first two HIPC assessments included 23 and 26 countries respectively. PEFA assessments have been carried out since 2005 in more than 50 countries. However, the number of countries for which all three reports (two HIPC AAP and one PEFA) are available is 15.⁹ Twelve countries are in Sub-Saharan Africa and three in Latin America.

Conclusions on the rate of PFM reform progress must therefore be taken with caution. Institutions take time to develop and function effectively, and detecting changes in system performance arising from improvements in any specific component take longer. The HIPC Assessments recommended a maximum assessment frequency of not more often than once every 2-3 years. PEFA has recommended repeated assessments to take place every 3-5 years. For the 15 countries presently under consideration, the date of the most recent PEFA assessment varies from country to country. For some countries the length of time between the

⁹ These are Benin, Burkina Faso, Ghana, Guinea, Guyana (for which not a PEFA report, but an updated HIPC assessment is available), Honduras, Madagascar, Malawi, Mali, Mozambique, Nicaragua, São Tomé and Príncipe, Tanzania, Uganda and Zambia.

2004 HIPC assessment and the PEFA assessment may have been little more than a year, while for others it is three years.¹⁰

Although the PEFA framework used the HIPC assessments as the base, differences in scope, coverage, indicator design, and country assessment report data result in only 11 of the 16 HIPC indicators being used for this update. Table 1 summarizes the process of mapping PEFA indicators and information back to the HIPC methodology. For each country, a table bringing together the relevant text from the PEFA reports was prepared to support country-level indicator scores. As an example of how this was done, the table for Ghana is included in the appendix.

Table 1. Mapping HIPC and PEFA indicators

	HIPC Indicator	PEFA	Notes
Formulation	1. Coverage of the budget or fiscal reporting entity	Intro PI-8 PI-9	The HIPC indicator is not directly related to any one PEFA indicator, but information is often available in different parts of the PEFA report.
	2. Degree of spending being funded by inadequately reported extra-budgetary sources	PI-7 (i)	Directly related. Conversion sometimes difficult due to differences in thresholds.
	3. Reliability of the budget as a guide to future outturn	PI-1 PI-2	Directly related and easily convertible.
	4. Inclusion of donor funds	PI-7 (ii)	Directly related and easily convertible.
	5. Classification	PI-5	Directly related and easily convertible.
	6. Identification of poverty-reducing spending		Information not available in PEFA reports. Indicator dropped.
	7. Integration of medium-term forecasts	PI-12	Directly related and easily convertible.
Execution	8. Evidence of budget execution problems – arrears	PI-4	Directly related. Limited information available in some PEFA reports.
	9. Effectiveness of internal control system	PI-18 PI-20 PI-21	Easily convertible based on the information available for different indicators.
	10. Tracking surveys are in use	PI-23	Directly related and easily convertible.
	11. Quality of fiscal information	PI-22 (i)	Directly related and easily convertible.

¹⁰ This paper considers the date of completion of the first draft PEFA report, right after the collection of the relevant updated information as the relevant date for the update, rather than the publication of the final report. This is in line with the previous HIPC reports, and allows for a more precise judgment of progress made.

Reporting	12. Regularity of timely internal fiscal reporting		Information not available in PEFA reports. Indicator dropped.
	13. Regular fiscal reports track poverty spending		Information not available in PEFA reports. Indicator dropped.
	14. Transactions are recorded in the accounts in a timely fashion		Information not available in PEFA reports. Indicator dropped.
	15. Timeliness of audited financial information	PI-26	Directly related and easily convertible.
Procurement	16. Efficiency and effectiveness of the public procurement system	PI-19	Indicator dropped because (a) not included in 2001 HIPC assessment, and (b) PEFA indicator looks at different issues.

The reduced number of original HIPC indicators means the three categories used to summarize the results --- budget formulation, execution, and reporting --- are less meaningful for this update. For budget reporting, only one original indicator is included in this update. For procurement, the differences between the HIPC and PEFA methodologies mean that meaningful comparison is difficult. As a consequence, in addition to presenting results along the original HIPC AAP formats, this paper also presents the results by three basic dimensions of budgeting:

- a) the **quality of budget information** (based on HIPC indicators 1, 2, 4 and 5), to assess the coverage, comprehensiveness and clarity of the budget documents;
- b) looking at **the budget as a reliable policy instrument** (based on HIPC indicators 3, 7 and 10), in order to check the extent to which budgets are implemented as approved, contain a medium-term perspective and allow for tracking of resources to service delivery points; and
- c) the **effectiveness of control and oversight functions** (based on HIPC indicators 8, 9, 11 and 15), to ensure that adequate mechanisms are in place to guarantee the respect of existing rules and procedures, and promote transparency and accountability.

Scoring and Aggregation Methods

HIPC and PEFA methodologies utilize different scoring systems. Firstly, HIPC scores range from ‘A’ (highest) to ‘C’ (lowest). For each indicator, a benchmark or minimal standard is also defined, and country results are reported on the number of benchmarks met. The PEFA methodology, on the other hand, has a wider scale (‘A’ to ‘D’, with the possibility of intermediate ‘+’ scores for composite indicators), and reports raw scores rather than defining benchmarks. In this paper, PEFA information and scores are used to replicate the HIPC ‘A’ to ‘C’ scale.¹¹ Results are summarized according to both the *number of*

¹¹ Given the larger informational needs of the PEFA methodology, the reverse would be much more difficult.

benchmarks met and improvements in *raw scores* (including assigning a numerical value to each score¹² and adding them up, either by country or by indicator).

Reporting on raw scores rather than benchmarks met will reveal more variation in performance. The benchmark approach, as a fixed standard, would be expected to show less overall variation over time. While useful for assessing progress towards those standards, it can mask performance changes more generally. Countries might improve or decline in performance on raw scores, but still meet or fail to meet a benchmark. As noted above, performance is reported relative to both benchmarks and raw scores, but going forward raw performance scores are preferable.

Other Issues and Limitations

Information on Action Plans and on donor activities is not updated in PEFA assessments. HIPC country reports included sections detailing an action plan indicating short- and medium-term actions planned by the country to redress weaknesses identified, and a table indicating current donor assistance in support of the action plan or more generally in the area of strengthening PFM systems. This information is not available in PEFA reports, and therefore is not updated in this paper.

Caution is warranted in interpreting the results. As with any analysis, some degree of judgment was used in scoring, especially given the need to retro-fit PEFA information to update the scores on the HIPC indicators. While the PEFA assessments included field work in country, not all are final reports. Some are in draft, and may be revised as they become final. The present exercise is desk-based, and has not included any country-based collection of information. Finally, the small sample of countries covered limits the reach and scope of the conclusions that can be drawn from this exercise.

Countries showing a decline in PFM system performance on these eleven indicators might have actually made substantial progress in other PFM areas not covered. The 11 indicators included in the analysis, as noted above, primarily cover budget formulation and execution, with only one indicator covering budget reporting, and no coverage of procurement issues. Moreover, they only cover a sub-set of the areas now assessed through the PEFA framework. Therefore, the eleven indicators cannot be considered to cover a sufficient number of aspects of the entire PFM system.

This paper uses several different approaches to analyze the data and draw conclusions. It examines the results for the pool of countries under consideration looking at specific indicators, benchmarks met and raw scores. It covers average scores by country, by stage of the budget cycle, and according to new categories (quality of budget information, budget policy, and control and oversight). Finally, it looks at trends in performance

¹² C=1, B=2 and A=3.

throughout the period considered . The results of all of these approaches are the basis for the conclusions drawn in the final section of the paper.

4. Results

Most PEFA assessments for the 15 countries included in this study were carried out in 2006. Only in three countries (Zambia, Uganda and Mozambique) did the assessments take place in 2005, while in three more (Benin, Burkina Faso and Guyana) they were carried out in 2007. In the rest of the paper, for ease of presentation, the update will be referred to as the ‘2006 update’. Results are presented firstly according to the methodology used for previous Board Papers (benchmarks met and phases of the budget process), and then using the alternative methodology tracking raw scores and looking at basic dimensions of budgeting (Quality of Information, Budget Policy, and Control and Oversight)

2006 Update: HIPC benchmarks met

Overall, there has been limited progress in PFM system performance looking at total benchmarks met for the pool of 15 countries over the period 2001 to 2006. Six countries declined in PFM performance and five improved over the period, but the total number of benchmarks met for all countries was greater than the sum of declines in benchmarks met. Between 2001 and 2004 the total number of benchmarks met remained unchanged at 68¹³, while it increased from 68 to 72 between 2004 and 2006, an increase of about 6 percent. In the same period, the average number of benchmark met by country also increased from 4.5 to 4.8 on a total of 11 indicators (see Table 2). This means that, on average, countries are still achieving less than half of all benchmarks. However, it also indicates that, for this reduced set of indicators and countries, the rate of progress in meeting benchmarks improved, with no progress in average benchmarks met over 2001-2004, and a slight improvement over a generally shorter time period from 2004 to 2006.

Table 2. Benchmarks met by Country, 2001-2006 (out of 11 benchmarks)

	HIPC	HIPC	PEFA		
	2001	2004	2005	2006	2007
Benin	6	6			3
Burkina Faso	6	7			8
Ghana	1	4		7	
Guinea	4	4		3	
Guyana	5	6			7
Honduras	5	3		5	

¹³ Based on only 11 rather than the original 15 indicators, these results are different from those reported in the 2005 Board Paper. While results for all 15 indicators showed an improvement between 2001 and 2004, this is not the case for the 11 indicators included in this analysis.

Madagascar	6	4		3	
Malawi	4	3		3	
Mali	7	7		6	
Mozambique	4	3	4		
Nicaragua	4	4		4	
S. Tome & Principe	2	2		0	
Tanzania	6	7		9	
Uganda	6	6	6		
Zambia	2	2	4		
Total	68	68		72	

Note: black cells indicate a negative trend in PFM systems quality, grey ones a positive one, and white cells an unclear trend.

Progress, however, is very uneven. Some countries have made more progress than others over the whole period, notably Burkina Faso, Ghana, Guyana, Tanzania, and Zambia (see Table 2). These five countries experienced increases of up to six benchmarks (Ghana, which started from the lowest score), or managed to improve even over a limited period of time (in the case of Zambia, two additional benchmarks were met in the period of one year, between 2004 and 2005). On the other hand, six countries experienced a decline, including Mali, which had the highest initial score. In Benin the number of benchmarks met halved between 2004 and 2007, while São Tomé and Príncipe's performance declined from 2 benchmarks in 2004 to no benchmarks met in 2006. In the four remaining countries, there is no clear trend. In Uganda and Nicaragua, the number of benchmarks met remained unchanged, while in Mozambique and Honduras it declined but then recuperated.

The group of better performing countries remains similar, but is expanding. In 2001, the best performing country was Mali, with 7 of the 11 benchmarks met. In 2004 the number of better performing countries had grown to three (Burkina Faso, Mali and Tanzania), although the top score had not changed. By 2006/7, there were four countries (Burkina Faso, Ghana, Guyana and Tanzania) meeting 7 or more benchmarks, while the maximum score increased to 9 (Tanzania).

Box 1. Burkina Faso: Long history of PFM reforms and better ownership as drivers of success?¹⁴

Burkina Faso has a long history of public financial management (PFM) reforms, pre-dating both the HIPC and PRSP initiatives and their related processes. The devaluation of the CFA franc in 1994 gave a strong impetus to macroeconomic and PFM reforms. The West African regional integration process was also an important factor, as it encouraged governments to undertake the reforms

¹⁴ This box is based on information drawn from the PEFA report of March 2007, the evaluation of General Budget Support available at <http://www.oecd.org/dataoecd/41/61/36685464.pdf>, and PREM Note 35 (<http://www1.worldbank.org/prem/PREMNotes/premnote35.pdf>).

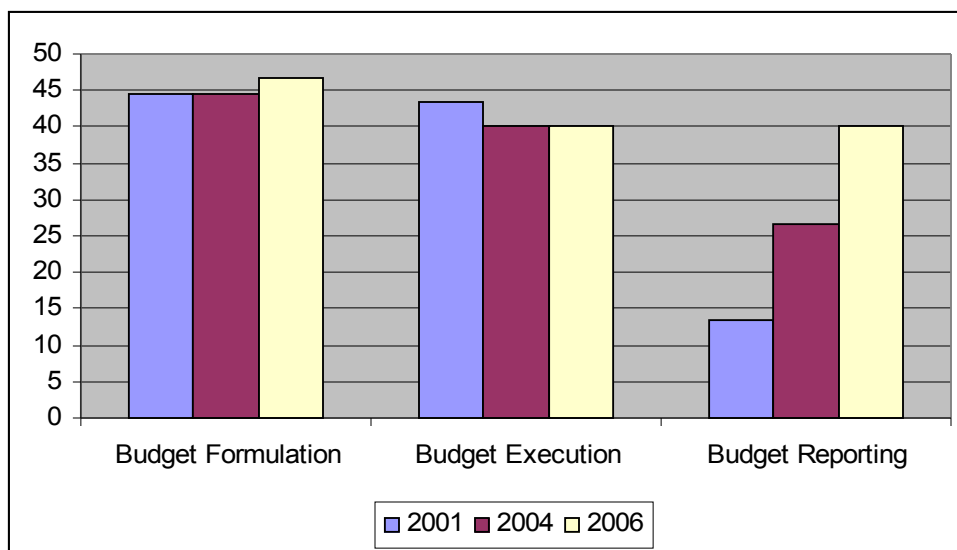
necessary to comply with the West African Economic and Monetary Union (WAEMU) convergence criteria and standards for public finance. The pilot initiative on new conditionality approaches led in the late 90s by the World Bank also played a key role in drawing attention to PFM issues.

Cabinet adopted the first comprehensive program to strengthen budget management in July 2002. Although it was originally devoted to expenditure reforms, since 2004 it has also encompassed revenue reforms. The *Programme de Reforme de la Gestion Budgetaire* (PRGB) built upon earlier reforms, and has been followed up by a series of initiatives, such as the *Stratégie de Renforcement des Finances Publiques* (SRFP) approved in February 2007. The main objectives of the SRFP are to improve the legal and operational framework for budget procedures, and to enhance, among other things, transparency and decentralization in budget management. The extensive dialogue and participatory process behind the formulation of these plans has led to a good degree of ownership and of responsibility of the main actors involved.

In Burkina Faso, centralized systems have resulted in orderly, if cumbersome, budget processes and execution. Computerization is fairly advanced and systems are able to produce various regular financial documents. Since 2003, there has also been a greater effort to prepare the final government accounts within the period allowed by the regulations. Measures to de-concentrate the budget have been initiated with a view to enhancing operational efficiency through timelier budget execution, although progress has been mixed.

Improvements in PFM system performance in terms of benchmarks met across the different phases of the budget cycle was also uneven, with budget formulation improving, execution slightly deteriorating, and too few indicators in reporting to draw a firm conclusion. Using the sub-division utilized in previous Board Papers, and looking at progress in the different phases of the budget cycle (formulation, execution and reporting) in terms of benchmarks met, positive changes can be seen in the areas of budget formulation and reporting (see Figure 1). The area of budget execution saw a slight deterioration in the 2001-2004 period, mostly due to the persistence of problems with expenditure arrears in a number of countries. The great improvement in the area of budget reporting is misleading as it is based on information for a single indicator (indicator 15 on timeliness of audited financial information), but reflects a very encouraging positive trend in budget accountability, with audit institutions and parliaments playing a larger role thanks to the regular and timely production of year-end accounts by the executive.

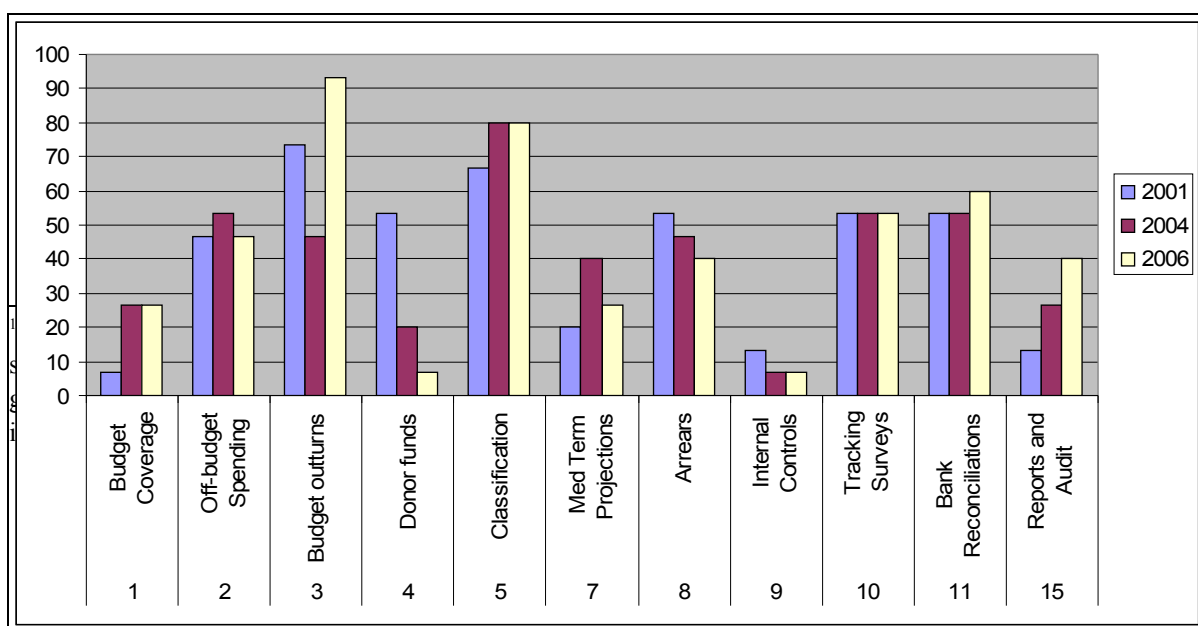
Figure 1. Percentage of Benchmarks met by Phase of budget cycle



Performance on benchmarks met for individual indicators is mixed. Figure 2 gives the details, which show that, by 2006:

- **93 percent of countries were able to limit the discrepancies between budget allocations and budget outturns** (indicator 3), compared to 73 percent in 2001 and 47 percent in 2004;
- **less than 10 percent of countries could fully reflect donor funds in their budget documents** (indicator 4), down from more than 50 percent in 2001¹⁵;
- **improvements in budget classification were sustained from 2004** (indicator 5), with 80 percent of countries meeting the benchmark;
- **there was a decline in the quality and integration of medium-term projections in budget processes for 2004** (indicator 7); and
- **budget coverage** (indicator 1) **and effectiveness of internal controls** (indicator 9) **remain problematic.**

Figure 2. Percentage of Countries Meeting Benchmark by Indicator



Countries showing improved performance tended to remain on an upward path.

In terms of benchmarks met, countries that improved between 2001 and 2004 tended to continue to improve between 2004 and 2006, while countries that showed no improvement or declined between 2001 and 2004 tended to continue to show no improvement or decline, though some did reverse course. More specifically:

- all countries that improved over 2001-2004 also improved over 2004-2006;
- about two-thirds of countries showing no change in performance over 2001-2004 declined over 2004-2006, nearly one-third remained constant, with one country reversing and showing improvement;
- one-half of countries with declining performance over 2001-2004 reversed course and showed improvement over 2004-2006, one-quarter remained constant and one-quarter continued to decline.

More research is needed to understand the dynamics of change, but these trends suggest that countries with positive reform momentum can maintain it over time (though 5-6 years is admittedly a short period of time), while countries showing declining performance can and do reverse course even within a short timeframe. Studying in more detail countries which managed to reverse course might be particularly useful in understanding reform dynamics.

2006 Update: Raw indicator scores

As noted previously, PFM system trends are more pronounced when examining raw indicator scores rather than the number of benchmarks met. The analysis based on ‘number of benchmarks met’ focuses on how close or far country PFM systems are from a set of performance benchmarks. Country PFM systems meeting these benchmarks would be deemed to be able to track and report reliably on pro-poor spending. But the benchmark-met approach masks changes in the underlying indicators. This section presents results of PFM system performance based on raw indicator score changes, using a numeric convention for ease of analysis and representation. The underlying HIPC A-B-C scoring method is

converted to numeric values 3-2-1, respectively. In terms of total possible raw scores for 11 indicators, therefore, a perfect score would be 33 points.

Overall, raw indicator scores show that there has been more widespread improvement in PFM system performance in the 15 countries assessed, although still at a modest pace. Using the different methodology that assigns a numerical value to each indicator, overall progress shows that total numerical scores increased from 321 in 2001, to 327 in 2004 and 332 in 2006¹⁶. This represents a modest increase over the whole period. The average country score also increased from 21.4 to 22.1, out of a possible maximum score of 33, roughly corresponding to a ‘B’ score.

The country-level analysis also indicates improvement, though with more variation. The total number of countries improving their PFM system performance between 2001 and 2006 increased to seven (now including Mali and Nicaragua), while countries experiencing declining PFM system performance fell to four (now including Uganda, see Table 3). Among improving countries, the average increase was of 3.5 points, with mixed performance across indicators. Ghana once again had the highest increase (10 points). For countries where performance worsened, the average fall was of 3.75 points, with Benin falling 5 points, and no country improving on any indicator.

Countries showing improved PFM system performance tend to continue to improve, while countries with declining performance tend to have a mixed performance (half stagnate or continue to decline, but half reverse course). In terms of raw indicator scores by country, most countries that improved over 2001-2004 continued to improve, and most countries that showed declining performance over 2001-2004 changed course to improve or hold constant over 2004-2006. More specifically:

- all countries showing no improvement in PFM system performance over 2001-2004 showed declining assessed system performance over 2004-2006;
- one-half of countries declining over 2001-2004 reversed course and improved (though in some cases only to 2001 levels) over 2004-2006, one-quarter showed no change, with the remaining quarter continuing to show declining performance;
- two-thirds of countries that improved over 2001-2004 continued to improve over 2004-2006, though twenty percent did reverse course and decline over 2004-2006.

Table 3. Overall progress by Country (total raw scores)

	HIPC	HIPC	PEFA		
	2001	2004	2005	2006	2007
Benin	24	24			19
Burkina Faso	24	25			26

¹⁶ Given that numerical values were assigned arbitrarily, relative differences in overall scores cannot be relied upon to change the rate of change.

Ghana	15	22		25	
Guinea	19	21		19	
Guyana	22	23			25
Honduras	23	20		23	
Madagascar	22	23		22	
Malawi	23	20		19	
Mali	25	26		26	
Mozambique	21	19	22		
Nicaragua	19	20		22	
S. Tome & Principe	18	18		14	
Tanzania	24	25		27	
Uganda	25	23	23		
Zambia	17	18	20		
Total	321	327		332	

Note: black cells indicate a negative trend in PFM systems quality, grey ones a positive one, and white cells an unclear trend.

In some cases, the different analytic reference point (benchmarks met versus raw scores) suggests different PFM system performance. For example, looking at total number of benchmarks met, Mali's performance remained constant between 2001 and 2004, and declined between 2004 to 2006. On the other hand, raw scores show an improvement between 2001 and 2004, and constant performance between 2004 and 2006. For Uganda, declining PFM system performance becomes more evident. These changes are related to the fact that countries may be making progress even without managing to increase the number of benchmarks met, or may have declining assessed PFM performance even though they manage to maintain the overall number of benchmarks.

Box 2. Malawi: PFM systems at a turning point?¹⁷

The PFM system in Malawi has been characterized by serious flaws in a broad range of areas. Budgeting, financial control, cash management, accounting and auditing all suffer from significant weaknesses. Areas where performance has been better have been in revenue management, where Malawi's performance reflects well against regional comparators, and in the use of pro-poor spending indicators.

In the past, the Government of Malawi has drawn up comprehensive sets of measures required to strengthen public financial management and accountability, for example as set out in the Malawi Financial Accountability Action Plan (MFAAP), finalized in 2003 and approved by Cabinet in May 2004. Nevertheless, only limited progress has yet been achieved in the implementation of the large number of tasks identified within it, largely due to its complexity and lack of focus. The Government of Malawi and donor efforts in a number of demanding fields – such as the

¹⁷ This box is based on information drawn from the PEFA report of September 2006 and the evaluation of General Budget Support available at <http://www.oecd.org/dataoecd/41/60/36685593.pdf>.

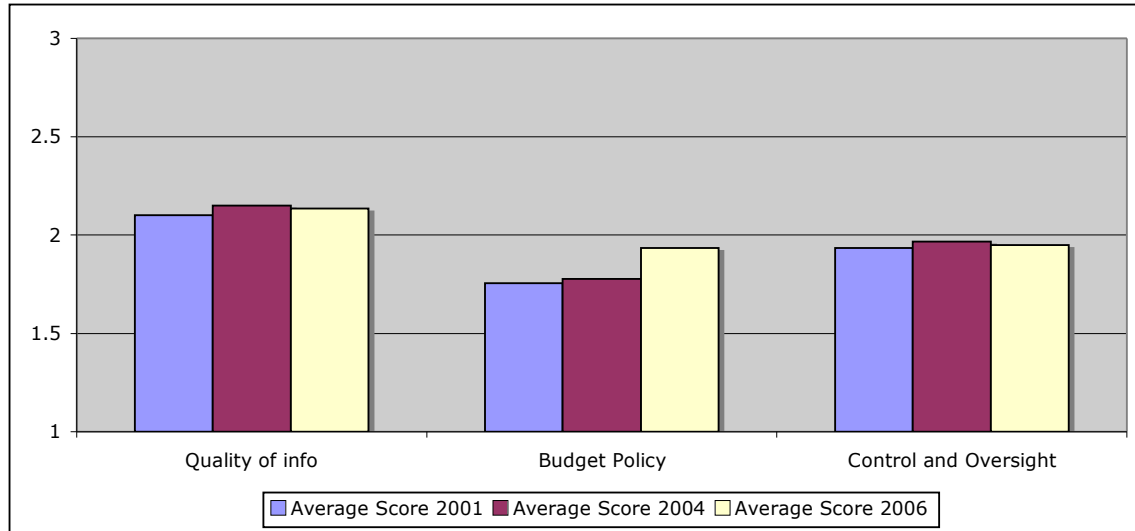
introduction of an Integrated Financial Management Information System (IFMIS) and a Medium Term Expenditure Framework (MTEF) – have been quite disappointing and have had only limited success. This is suggestive of a lack of government commitment, or lack of broader consensus within government, to effective PFM systems. The lack of commitment is compounded by the fact that donors have encouraged the Government to adopt complicated budget management systems such as Output Based Budgeting in an environment of limited capacity.

Some improvements have been made since 2004 by the new government, mostly in the exercise of fiscal discipline and fiscal control. This has enabled some of the foundations laid in the earlier period to play a more productive role. Standard operating procedures for government processes seem to have been allowed to operate with less resort to exceptional procedures or *ad hoc* interventions. Additionally, there are signs that the parliament is taking its role in PFM more seriously, with improved scrutiny of executive action. However, concerns remain over the level of compliance by Ministries with existing PFM regulations, the continuing lack of enforcement of some sanctions, and the delays in certification of accounts by the National Audit Office.

The analysis of overall progress by budget dimension reveals a more mixed picture.¹⁸ The average scores for the quality of budget information and for the effectiveness of control and oversight functions remain largely unchanged, with slight deterioration in the period 2004-2006. The only clear improvement is in the degree to which budgets are a reliable policy instrument, driven by the increase in average scores on budget outturns (indicator 3).

Figure 3. Average country scores by budget dimension

¹⁸ Note: as with benchmarks, reporting results in aggregated formats reduces variation and masks some trends.

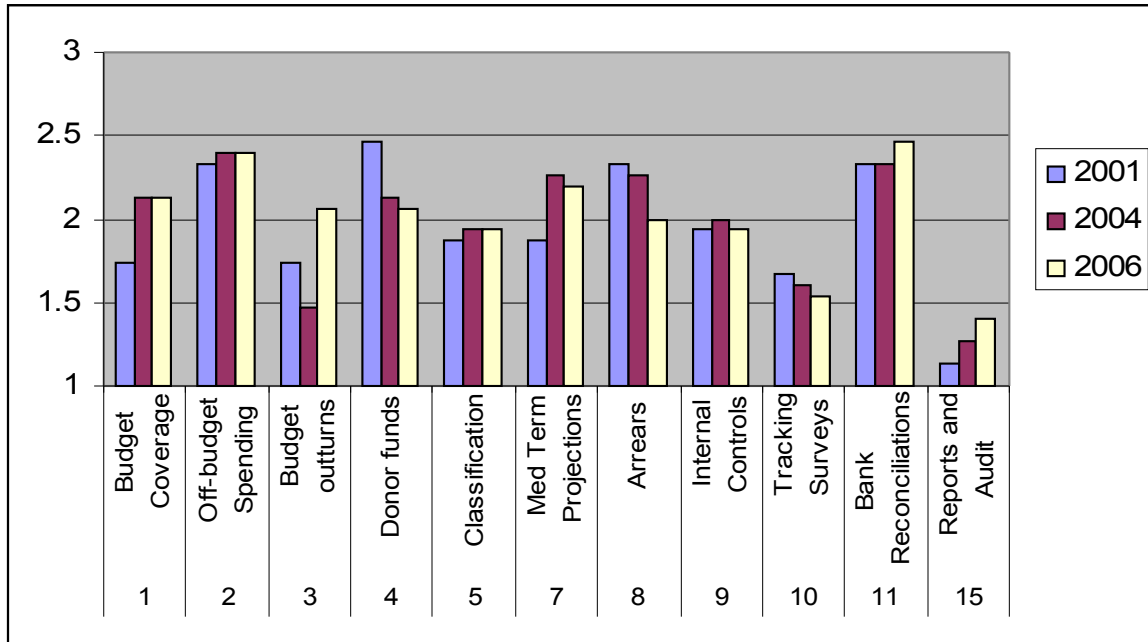


Looking at performance by indicator, the results show similar patterns for raw score changes as for benchmarks met. Average scores by indicator (see Figure 4) show that performance changes less dramatically than for benchmarks met, but follows similar patterns. In particular:

- Indicators with the best average score for 2006 are the ones on **bank reconciliations** (indicator 11), on **off-budget spending** (indicator 2), and on **medium-term forecasts** (indicator 7);
- Indicators with the lowest average scores in 2006 are the ones on **external audit** (indicator 15) and on the **use of tracking surveys** (indicator 10);
- There are clear improvements on five indicators over the whole period 2001-2006, a worsening on three of them, and an unclear trend for the remaining three;
- The greatest improvements have been seen for **budget coverage** (indicator 1) and **external audit** (indicator 15), while the worst decline is associated with **inclusion of donor funds**¹⁹ (indicator 4) and **expenditure arrears** (indicator 8).

Figure 4. Average scores by Indicator

¹⁹ See the caveat in footnote 13.



While countries' performance has seen limited change in each of the budgeting dimensions, the cumulative effect has been in some cases notable. Looking at the combination of results by country and by dimension, Table 4 shows that the majority of countries, in each of the budgeting dimensions, have seen limited change over time. However, the cumulative effect of such changes sees a majority of countries witness a clear although limited improvement in their overall score. For five countries, overall score has increased by 3 or more points. At the same time, three countries have seen a decline of the same magnitude.

Table 4. Country changes in Overall Scores, by Dimension and Total

Change 2001-2006	Quality of Info	Budget Policy	Control & oversight	TOTAL
>=5	-	-	Ghana	Ghana
3/4	Ghana	Zambia	-	Guyana,

				Nicaragua, Tanzania, Zambia
1/2	Burkina Faso, Guyana, Mali, Mozambique, Nicaragua, Tanzania	Ghana, Guinea, Guyana, Mali, Mozambique, Nicaragua	Burkina Faso, Honduras, Tanzania, Zambia	Burkina Faso, Mali, Mozambique
0	Guinea, Madagascar	Honduras, Madagascar, São Tomé and Príncipe, Tanzania, Uganda	Guyana, Madagascar, Nicaragua, Uganda	Guinea, Honduras, Madagascar
-1/2	Benin, Honduras, Malawi, São Tomé and Príncipe, Uganda, Zambia	Benin, Burkina Faso, Malawi	Benin, Guinea, Malawi, Mali, Mozambique	Uganda
-3/4	-	-	São Tomé and Príncipe	Malawi, São Tomé and Príncipe
<=-5	-	-	-	Benin

5. Conclusions and Suggestions for Future Research

The results presented in this paper need to be interpreted cautiously. Only 15 out of 26 countries and 11 out of 16 indicators included in previous HIPC assessments were used for the entire period 2001-2006. The limited sample size of countries and indicators means that the trends may not be robust or accurately reflect country changes. Effectively, the indicators used look at only a small part of country's PFM systems. In some cases, the mapping of the information back to the HIPC methodology was not straightforward, and a certain degree of judgment was necessary in providing updated HIPC scores with PEFA information. As a consequence, a series of adjustments to the scoring and aggregation methodologies were made in order to allow for a more meaningful assessment on the basis of the information available.

The emerging overall trend in quality of PFM systems over the whole 2001-2006 period for the 15 countries under analysis is one of limited and uneven progress, but with about half of the countries making steady improvements. The number of countries performing well has increased, although performance has varied substantially. Therefore, having realistic expectations of the possible rate of PFM system improvement is critical. For donors, this suggests that longer-term, sustained commitments to PFM system reform may be a critical element of success. This is consistent with understanding of organizational development and institution-building requirements.

Clearly, there is a need to better understand the underlying dynamics leading to an improvement in performance. The information summarized in this paper can be used for a number of purposes. Some of the most interesting ones are related to follow-up research that can shed some light on the underlying factors and dynamics of country-level progress (or lack thereof), both overall and in specific dimensions and indicators. Box 3 presents the summary of a preliminary attempt that deserves to be built on.

Box 3. Explaining PFM system improvements using 2001-2004 HIPC results²⁰

Many factors may influence PFM system performance. Some of the main ones include: (1) initial country PFM system performance; (2) government commitment to reform; and (3) donor assistance supporting reform and capacity-building. Generally, lower initial PFM capacity and lower government commitment to reform suggest a slower pace of reform, while more donor assistance is presumed to lead to greater progress.

A model was tested using data from the 2001-2004 HIPC assessments. Initial PFM system performance was defined as the number of benchmarks met in 2001; government commitment as the number of actions fully implemented and initiated; and donor involvement as the number of donor agencies actively providing assistance. The dependent variable – change in PFM system performance – was defined as the change in number of benchmarks met between 2001 and 2004.

Preliminary results indicated that initial country PFM system conditions and government commitment to reform were significantly correlated with improvements in PFM system performance. Donor support was not found to be significantly correlated with improved PFM system performance.

This analysis could be refined in a number of ways, accounting for possible omitted factors, and using more other measures of government commitment or of donor support, such as for example the total amount of financial resources invested by donors on supporting PFM reforms over a certain period.

In the future, similar updates should be done on the basis of PEFA assessments and methodology. This paper follows up on previous Bank-Fund Board Papers in a transitional phase where HIPC assessments have been discontinued and the more elaborate PEFA methodology is being implemented. PEFA assessments represent a more detailed instrument for assessing PFM system quality, and they are being carried out in a much wider set of countries, with the intention of repeating them over time at 3-5 year intervals. This will generate a much improved set of data and information, covering more PFM areas (such as, for example, taxation and transparency) for more than just HIPC countries. In subsequent papers, therefore, the PEFA database should be used as the basis for comparative assessments, with previous HIPC indicators simply being looked at to assess the degree of past performance and progress in certain areas, for only a subset of countries, over a longer period of time.

²⁰ Greg Kisunko and Shilpa Pradhan (WB) contributed substantially to this analysis.

Table 5. Overall Results by Country and by Indicator

Country	HIPC Indicator (2001, 2004, 2006)																																	
	1			2			3			4			5			7			8			9			10			11			15			
Benin	B	B	B	A	A	C	B	B	B	B	B	B	B	B	B	B	A	B	A	B	C	B	B	B	B	B	C	A	A	A	C	C	C	
Burkina Faso	B	B	A	A	A	A	B	B	B	C	C	B	B	B	B	A	A	A	B	A	A	B	B	B	A	A	B	A	A	A	C	C	C	
Ghana	C	B	B	B	B	A	C	C	B	B	B	C	B	B	B	B	C	A	A	C	B	B	B	B	B	B	C	A	A	C	C	B		
Guinea	C	A	B	B	B	B	B	C	B	A	B	B	B	B	B	C	B	B	B	C	C	C	B	C	C	C	C	A	A	A	C	B	C	
Guyana	A	B	A	B	A	A	B	B	B	A	A	A	C	B	B	B	B	B	A	A	A	B	B	B	C	C	B	C	C	B	B	B	C	
Honduras	B	B	B	A	A	A	B	C	B	A	B	B	A	B	B	B	B	B	A	A	B	B	B	B	C	C	C	C	C	A	C	C	B	
Madagascar	B	B	B	A	A	A	B	C	B	A	B	B	B	A	A	C	A	B	C	B	B	B	B	B	B	B	B	C	A	B	B	C	C	C
Malawi	C	C	C	A	A	A	C	C	B	B	B	B	A	B	B	B	B	B	A	B	C	B	B	B	A	B	C	B	B	B	C	C	C	
Mali	B	A	A	A	B	B	B	B	A	B	B	B	B	A	A	B	A	A	A	A	B	A	B	B	B	B	B	A	A	A	C	C	C	
Mozambique	B	A	A	C	C	B	B	B	A	B	B	B	B	C	C	B	B	B	A	A	A	A	B	B	C	C	C	B	C	C	C	C	B	
Nicaragua	C	A	B	C	C	B	B	B	B	A	A	B	B	B	B	C	C	B	B	B	A	B	B	B	C	C	B	A	B	B	C	C	C	
S. Tome & Principe	B	B	B	B	A	B	C	C	C	A	A	B	C	C	C	C	C	C	A	B	C	C	C	C	C	C	C	B	B	C	C	C	C	
Tanzania	B	B	B	A	A	A	B	B	B	B	B	B	C	B	B	A	A	A	A	A	A	B	B	A	B	B	B	B	A	A	A	C	C	B
Uganda	B	B	B	B	B	B	B	C	B	A	B	B	A	B	B	A	A	A	B	C	C	B	A	B	B	B	B	B	A	A	B	B	B	
Zambia	C	C	C	B	B	B	C	C	B	A	B	B	C	C	C	C	B	B	C	C	C	B	B	B	C	C	B	A	A	A	C	B	B	
Benchmark	A			A			B			A			B			A			A			A			B			A			B			

Note: For each indicator, the three boxes correspond to the 2001, 2004 and the most recent PEFA assessments. Letters in bold denote benchmarks met.

Appendix

Country Table for Ghana (First draft: May 2006, Final report: June 2006)

HIPC Indicator	PEFA	Relevant text from PEFA assessment	2001	2004	2006
1. Coverage of the budget or fiscal reporting entity	PI-8	There is no consolidation of central and sub-national fiscal data for the general government sector, as the two levels use different classification systems.	C	B	B
	PI-9	The State Enterprises Commission receives reports from the majority of SOEs, at least annually, but does not prepare a consolidated report			
2. Degree of spending being funded by inadequately reported extra-budgetary sources	PI-7 (i)	<p>Significant progress has been made during the last year in the comprehensiveness of reporting on domestic fiscal operations. In particular, the capture of non-tax revenues has improved substantially with the reporting of the collection and use of all internally-generated funds (IGFs) (both lodged and for the first time also the retained part) by MDA in the 2005 and 2006 Budget Statements.</p> <p>Annexes in the 2005 and 2006 Budget Statements contained the income and expenditure statements for the last completed financial year (for 2004 in the case of the 2006 Budget) for the four Statutory Funds operating at that time.</p> <p>Whilst the fiscal reports do not include information on Government guarantees, these are relatively limited (representing around 3 percent of expenditures), and information on them is available from MoFEP</p>	B	B	A
3. Reliability of the budget as a guide to future outturn	PI-1	<p>Based on domestically-financed primary expenditure, deviations between original budget and outturns were (according to the administrative classification):</p> <p>2003 – 2.3% 2004 – 12.4% 2005 – 8.0%</p> <p>Deviations between original budget and outturns were (according to the</p>	C	C	B

	PI-2	administrative classification): 2003 – 13.6% 2004 – 13.4% 2005 – 25.3%			
4. Inclusion of donor funds	PI-7 (ii)	Since donor funds are only partially reported (grants included in the CAGD report represent about half of the amount estimated by ADMU) together with CF funds, retained IGF resources, and HIPC resources, there is not a single source for reporting on total Government fiscal operations. This hampers effective oversight of the use of public resources. Available data from ADMU indicate that around 94 percent of income and expenditure on donor-funded projects are captured in fiscal reports. Capture of donor flows in fiscal reports has increased in recent years.	B	B	B
5. Classification	PI-5	The budget documents and in-year and end-year fiscal reports are presented annually on the basis of the administrative classification and by aggregated economic item (personal emoluments, non-salary administration, services, and investment). In addition, since the current Chart of Accounts is not fully consistent with the GFS, a bridge table, converting the budget by MDA to the 10 main COFOG functions, has been prepared for each of the last three budgets. This functional table is included in the individual volumes for MDA Estimates but is not shown in the Budget Statement.	C	B	B
7. Integration of medium-term forecasts	PI-12	A medium-term expenditure framework (MTEF) has been in place since the PFM reforms that were launched in the mid-1990s. Currently, the budget is set within the context of the MTEF, which provides the aggregate fiscal framework on a rolling three-year basis, and determines the overall resource envelope for the medium term, as well as the indicative discretionary resource allocations amongst MDAs. Nonetheless, there is concern about the operationality of the MTEF, with an apparent disconnect between the MTEF and budget implementation, as evidenced by the significant variances between budget plans and outturns discussed above.	B	B	B
8. Evidence of budget	PI-4	In 2005, total arrears (roads plus non-roads) were 131 bn cedis,	C	A	A

execution problems – arrears		representing 1.1% of total GoG discretionary expenditure.			
9. Effectiveness of internal control system	PI-18	Full effectiveness of the payroll and personnel system has been hampered due to technical problems, capacity constraints in managing the system, and weaknesses in security controls.	C	B	B
	PI-20	The GoG has improved its control of commitments with the introduction of a new expenditure commitment control system (CCS) in September 2003. Whilst the new commitment control system has improved expenditure control for the majority of expenditures, a limited amount of expenditures remain outside of the CCS.			
	PI-21	In general, internal rules and procedures appear to be understood by those directly and routinely involved but clear understanding of all of the rules is not necessarily widespread (e.g. amongst all of those with signatory responsibility) Insufficient understanding of the role of internal audit in ensuring appropriate control in an environment where the Accounting Officer is formally held accountable for his/her expenditure, combined with weak capacities amongst IAU staff in the MDAs, undermine the effectiveness of the internal audit function in carrying out these roles.			
10. Tracking surveys are in use	PI-23	No comprehensive surveys of resources (cash and in-kind) going down to the level of schools and/or health centres have been conducted during the last three years, although plans are currently being developed to undertake one in the education sector.	B	B	B
11. Quality of fiscal information	PI-22 (i)	Bank reconciliation takes place for all MDAs on a monthly basis and is completed within 8 weeks of the end of the month.	C	A	A
15. Timeliness of audited financial information	PI-25	The Financial Administration Act stipulates that Annual Financial Statements must be submitted to the Auditor-General within three	C	C	B

	PI-26	<p>months. This requirement has been met in recent years.</p> <p>There was a backlog of submissions of audit reports to Parliament, primarily of SOEs and SFs, which was largely cleared in 2005 with the help of contracted private auditors. The 2004 Auditor-General's reports were submitted to Parliament in October (for MDAs) and December (for CF) 2005, which was within the 12-month range.</p>			
--	-------	---	--	--	--