THIRD BENCHMARKING REVIEW

OF ECG MEMBERS' EVALUATION PRACTICES
FOR THEIR PRIVATE SECTOR INVESTMENT OPERATIONS
AGAINST THEIR AGREED GOOD PRACTICE STANDARDS

August 25, 2010
(Corrected September 22, 2010)

Walter I. Cohn & Associates, LLC
6212 Goodview Street, Bethesda, Maryland, 20817, USA
Telephone: (301) 229-5023 Fax: (301) 320-2529 E-mail: widacohn@comcast.net

CONTENTS

ABBREVIATIONS & DEFINITIONS

AfDB	African Development Bank
AsDB	Asian Development Bank
Best practice standards	Practices going beyond good practice standards that are desirable but not essential
EBRD	European Bank for Reconstruction & Development
ECG	Evaluation Cooperation Group (of the MDBs)
EIB	European Investment Bank
Experimental standards	Third Edition standards that were to be tried on an experimental basis with a view to deciding in 2010 whether to retain them, revise them, or abandon them.
Good practice standards	The key principles and practices that any development institution that finances the private sector should follow if it is to have a satisfactory evaluation system
GPS or GPS-IO	Good Practice Standards for Private Sector Investment Operations
Harmonization standards	Minimum standards necessary to permit comparability of reported operational results among the MDBs, as called for by the Development Committee
IADB	Inter-American Development Bank
IIC	Inter-American Investment Corporation
IFC	International Finance Corporation
IMF	International Monetary Fund
MC	Materially consistent with the substance of the GPS
MDB	Multilateral Development Bank
MIF	Multilateral Investment Fund, a part of the IADB Group
MIGA	Multilateral Investment Guarantee Agency
NC	Not materially consistent with the substance of the GPS.
NR	Not relevant because (a) The MDB cannot follow the standard because of (i) the nature of its operations (not its policies or practices) or (ii) its total reliance on evaluations carried out by the central evaluation department (rather than evaluations by the MDB's operational staff), or (b) the MDB meets an equivalent or relevant higher standard specified in the GPS.
NC	Not materially consistent with GPS standard
Other standards	Standards (other than *harmonization* standards) specified in ECG-GPS, 3rd edition, April 2006, not essential for comparability of reported results but desirable for enhanced evaluation relevance, accountability and learning within each institution
SCF	IADB's Structured & Corporate Financing Department, which is responsible for IADB's non-sovereign lending.
WBG	World Bank Group
WGPSE or the Working Group	The ECG's Working Group on Private Sector Evaluation

THIRD BENCHMARKING REVIEW
OF ECG MEMBERS' EVALUATION PRACTICES FOR PRIVATE SECTOR INVESTMEMENT OPERATIONS AGAINST THEIR AGREED GOOD PRACTICE STANDARDS

EXECUTIVE SUMMARY

i. Fourteen years ago, a Development Committee Task Force called for harmonization of evaluation methodologies, performance indicators and rating criteria by the Multilateral Development Banks (MDBs). Pursuant to this mandate, the MDBs' Working Group on Private Sector Evaluation (WGPSE) issued good practice standards for evaluation of private sector investment operations (GPS) in 2001 and then twice revised these standards. The present paper reports on the third independent assessment of the progress made.

ii. Nearly all the organizations covered have made significant progress towards harmonization. On average, their practices are now materially consistent with 69% of the good practice harmonization standards, well above the 59% for the good practice standards in 2004 and 39% in 2002. The second benchmarking report--which had been circulated to the boards of directors of most members--stimulated members to adopt more of the good practice standards, bring more of their practices into line with the standards, or both. In addition, the advice provided in the context of the current benchmarking review helped several members to correct shortfalls in their evaluation mandates or guidelines.

iii. Nevertheless, the members have not yet fully carried out the Development Committee's mandate. Indeed, some members' scores continue to fall well below an acceptable level. Moreover, even though most of the members use the GPS terminology to describe their rating categories, the content of these categories frequently differs. The members may use the same words for their rating dimensions and indicators, but the meaning of the words often differs from one institution to another. And despite the progress on benchmarks and methodological issues, the members have not yet achieved consensus on the experimental standards, introduced in the Third Edition to permit comparable assessments of contributions to business success and economic development, two of the four indicators for rating development outcomes.

iv. Two factors account for most of the shortfalls. First, some member institutions or their evaluation departments disagree with some standards. They have not accepted the decisions made by the WGPSE when it approved the standards. Second, some have not adequately committed themselves to adoption of the standards. Change requires time and effort. Some evaluation departments may have been unwilling to undertake the efforts needed to convince others within their institutions (possibly including their Boards) to adopt new standards, to insist on the need for using different standards for evaluating public and private sector operations, or to retrofit earlier evaluation ratings to permit internal comparability over time.

v. The members need to pursue their efforts towards harmonization, not only to fulfill the mandate laid down by the owners but also--and equally important--to ensure that their evaluation work is consistent with practices that will yield the most reliable, most useful results. To this end,

- *The next edition of the GPS should build on the progress already made and further clarify and strengthen the standards. It should not accept practices that can lead to misleading ratings on development outcomes, the MDB's investment profitability, the MDB's additionality, or the MDB's work quality; misleading judgments on overall results; and unjustified conclusions for future operations.*

- *The WGPSE should allocate sufficient time to reach consensus on the next edition of the GPS, including revisions to the experimental standards. Before the WGPSE meets to*

discuss the draft Fourth Edition, it should provide a period of several months for written comments and exchanges of views. Expecting to reach consensus in a single meeting, lasting less than a day, has proven to be unrealistic.

- *To facilitate consensus on performance benchmarks, the good practice standards should be limited to the benchmarks for "satisfactory" ratings and should treat the other benchmarks as best practice standards.*

- *The members who are lagging should explicitly commit themselves to bringing their mandates, guidelines and practices into line with the GPS by the next benchmarking review. They may be subject to some constraints that appear insuperable but should seriously consider what might be done to overcome these constraints.*

- *The next benchmarking review should allow sufficient time to assess whether a reasonably sized random sample of each member's evaluation reports is consistent with the GPS, i.e., whether the members are correctly applying the standards in practice. Correct application of the standards will lead not only to greater harmonization but--equally important--to evaluations that are more solidly grounded and, hence, more useful.*

THIRD BENCHMARKING REVIEW
OF ECG MEMBERS' EVALUATION PRACTICES
FOR PRIVATE SECTOR INVESTMEMENT OPERATIONS
AGAINST THEIR AGREED GOOD PRACTICE STANDARDS

1. BACKGROUND

Context

1. Fourteen years ago, a Development Committee Task Force called for harmonization of evaluation methodologies, performance indicators and criteria by the Multilateral Development Banks (MDBs):

> The development of objective indicators of performance is…essential for the public accountability of the MDBs and their ability to justify their use of public resources to shareholder governments, parliaments, and the public. Currently, it is not possible to compare their operational results, or even to describe them in a common language. Major public sector institutions like the MDBs must be able to account for their efforts in readily understood terms. A common methodology for evaluating their portfolios should be developed.... A determined effort should be made to harmonize performance indicators and evaluation criteria, taking into account the differing circumstances of each institution. The lessons learned from these evaluations should be shared among the MDBs with a view to applying them quickly in new operations.[1]

2. The MDB presidents endorsed these recommendations and called for "further intensification of collaboration among MDB evaluation units in harmonizing evaluation standards…."[2]

3. In February 2001, in response to this mandate, the Working Group on Private Sector Evaluation (WGPSE, or the Working Group) agreed[3] on a set of good-practice standards[4] for evaluation of private sector

---

[1] Development Committee, Task Force on Multilateral Development Banks, "Serving a Changing World—Report of the Task Force on Multilateral Development Banks," March 15, 1996, p. 18.

[2] Development Committee, "Report from the Multilateral Development Banks on Implementation of the Major Recommendations of the MDB Task Force Report," March 26, 1998, p. 4.

[3] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA's evaluation unit IEG-MIGA had not been created at that time, and therefore did not contribute to the development of the GPS. The relevance and applicability of the standards designed for MDBs' private sector *investment* operations to the operations of an insurance agency such as MIGA which underwrites political risk guarantees was therefore not considered at the time, since their application to MIGA could not be anticipated. This remained the case in later years also since MIGA was not part of the WGPSE's 2006 consensus-based development of the GPS3. To date the relevance and applicability of the GPS3' specific standards, definitions and benchmarks to an insurance agency which does not engage in "investment operations", but rather engages in "issuing [political risk] insurance" have not been explicitly considered by the WGPSE. Thus the Development Committee's mandate to *"take into account [MIGA's] differing circumstances"* is still a work in progress.

MIGA may not have contributed to the First Edition of the GPS, but MIGA has been covered by all three benchmarking exercises and did participate in the March 2006 WGPSE meeting that approved the Third Edition. Indeed, the Third Edition (para. 9) specifically states that certain standards do not apply to MIGA.

[4] MIGA's evaluation unit has requested that the following footnote be included in the report:

> Given that the GPS3 harmonization effort has not explicitly considered how the GPS apply to MIGA, ***special care must be taken when rating MIGA to ensure that a given standard, benchmark or definition truly applies to MIGA and that the assessment has taken account any differing circumstances of MIGA's insurance business.*** MIGA disagrees with a number of the **"NC"** ratings in the report on the grounds that the standard is not applicable to MIGA and the rating did not take into account the differing circumstances and nature of MIGA's insurance business, and holds they should have been rated **"NR".** Footnotes indicate the ratings that MIGA disputes on the basis of their applicability to the insurance business.

investment operations (the GPS-IO or, more briefly, the GPS).[5] As part of these standards, the members of the WGPSE agreed that they would "arrange for independent periodic crosscutting assessments of the extent to which these…good-practice standards are being applied in each member agency's evaluations and annual reporting, and report the findings to the MDB Presidents." In the Third Edition of the GPS, the WGPSE members agreed further that they would submit these periodic reviews to their boards of directors.

4. At the request of the WGPSE, I assessed the adoption and application of the GPS in 2002 and 2004 and again in 2009-10.[6,7] The current assessment covers the World Bank Group's IFC and MIGA; the IADB Group's Structured and Corporate Finance Department (SCF), Multilateral Investment Fund (MIF), and IIC; and the private sector operations of AfDB, AsDB and EBRD.[8]

5. Based on the earlier assessments, the Working Group agreed on a Second Edition of the standards in May 2003 and a Third Edition in April 2006.[9] The Third Edition, which introduced significant changes, distinguished between several types of standards:[10]

- **Harmonization and other standards.** Some of the standards are necessary to permit comparability of operational results among the MDBs, as requested by the Development Committee and supported by the MDB presidents. These are categorized as *harmonization standards*. Others are not needed for comparability but are nonetheless designed (as are the harmonization standards) to help improve accountability and learning within each institution. These are designated as *other standards*. Both types of standards are important.

- **Good-practice and best-practice standards.** The *good practice standards* lay down the key principles that any development institution that finances the private sector should follow if it is to have a satisfactory evaluation system. The *best practice standards* set out more detailed practices that are desirable but, because of requiring additional work, not essential.

See my comments on MIGA's notes elsewhere in this report and in Annex 1.

[5] MDB-ECG, WGPSE, "MDB-ECG Good-Practice Standards for Evaluation of Private Sector Investment Operations," April 23, 2001.

[6] Walter I. Cohn & Associates, LLC, "Benchmarking of ECG Members' Evaluation Practices for Their Private Sector Investment Operations Against Their Agreed Good-Practice Standards, October 18, 2002, and "Second Benchmarking Review of ECG Members' Evaluation Practices for their Private Sector Investment Operations Against Their Agreed Good Practice Standards," January 25, 2005.

[7] Terms of Reference, Third Assessment of WGPSE Member Practices against MDB-ECG Good Practice Standards for Evaluation of Private Sector Investment Operations, June 19, 2009, and Additional Terms of Reference, Third Assessment of WGPSE Member Practices against MDB-ECG Good-Practice Standards For Evaluation of Private Sector Investment Operations, January 15, 2010.

[8] The members of the WGPSE comprise AfDB, AsDB, EBRD, EIB, IFC, IADB, IIC and MIGA. The members of the ECG comprise AfDB, AsDB, EBRD, EIB, IADB Group, IMF, and WBG. The GPS are intended to apply to those investment or guarantee operations of the ECG members in developing and transition countries where there is no sovereign recourse for the MDB. Thus, the GPS apply to all the operations of the World Bank Group's IFC and MIGA and the IADB Group's Structured and Corporate Finance Department, IIC, and MIF, the bulk of the operations of EBRD, and smaller shares of the operations of AfDB and AsDB. To facilitate the exposition, "members" refers to the members of the WGPSE, rather than the ECG. EIB, which devotes roughly 85% of its resources to the European Union and does not fall within the purview of the 1996 harmonization mandate (despite accounting for a significant share of the MDBs' private sector commitments in developing countries), is not covered by this report.

[9] MIGA's evaluation unit has requested that the following footnote be included in the report: "MIGA was not included in WGPSE's 2006 consensus-based development of the GPS3. See footnote 3." See also my comment on footnote 3.

[10] ECG, WGPSE, "MDB-ECG Good-Practice Standards for Evaluation of Private Sector Investment Operations, Third Edition," April 18, 2006. Annex 2 of the current report sets down the criteria I used in interpreting the standards for this third benchmarking review as well as suggestions for possible changes in the standards.

- **Experimental standards.** The Third Edition designates several standards as experimental. These standards specify proposed benchmarks for rating contributions to business success, economic development and private sector development; lay out new approaches for cost-benefit assessments where normal financial and economic returns cannot be calculated; and provide a more concrete methodology for assessing the contributions to economic development of credit lines and investments in equity funds. The members were to experiment with these standards and to decide in 2010 whether to retain them, revise them, or abandon them.

- **Standards not universally applicable.** The GPS, Third Edition, recognizes that some standards are not applicable to certain MDBs because of the nature of their operations. For example, since some institutions do not make equity investments (or loans), standards relating to such operations would not be relevant for them.[11] Similarly, since MIGA's Independent Evaluation Group has so far carried out all evaluations itself, rather than calling for self-evaluation reports by the operational staff, standards relating to self-evaluation reports are not yet applicable to MIGA. And, since EBRD seeks to foster transition, rather than development and poverty reduction, EBRD has indicated that the standards relating to business success and economic development are not applicable to its operations.

6. The GPS, Third Edition, consists of 81 numbered standards. Some numbered standards, however, comprise a good practice standard and one or more best practice standards, so the total number of standards is 103. See Table 1 on next page.

7. This paper focuses on the 54 good practice harmonization standards, i.e., the minimum standards necessary to permit comparability of reported operational results among the MDBs. Many of these standards are substantially the same as in the Second Edition, but others are new or involve material changes. Of the 54 good practice harmonization standards, 37 are substantially the same as earlier, and 17 are new or involve material changes. The report includes a comparison of the results of the second and third benchmarking reviews for the 37 comparable good practice harmonization standards.

8. The remainder of the present chapter summarizes the methodology used for assessing the extent of harmonization. Chapter 2 reports on the member institutions' progress towards harmonization. Chapter 3 summarizes the findings with respect to the experimental standards introduced in 2006. Chapter 4 comments on the comparability between the rating categories used by individual organizations and the evaluative dimensions and indicators specified in the GPS. And Chapter 5 provides some concluding remarks and recommendations for moving further towards harmonization.

[11] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA does not make loans or equity investments. Thus standards relating to such operations would not necessarily be relevant for MIGA." As noted in an earlier footnote, the relevance and applicability of the standards designed for MDBs' private sector loan and equity operations to the operations of an insurance agency such as MIGA which underwrites political risk guarantees applicability to MIGA has not been reviewed to date by the WGPSE.

See footnote 3.

Table 1: Number of Standards, by Category

	2d Ed.	3d. Ed.	Comments on 3d Edition
Harmonization standards			
• Good practice standards	50	54	Including 3 with experimental components
• Experimental standards	0	10	Including 3 with non-experimental components
• Transitional standard	0	1	Optional standard for institutions with small volume of private sector investments
• Best practice standards	11	10	Several harmonization standards have one or more best practice options.
• Total harmonization standards	61	72	Net of 3 standards double-counted because of having both experimental & non-experimental components
Other standards			
• Good practice standards	21	19	Several "other" standards have one or more best practice options in addition to a good practice option; one "other" standard has only a best practice option.
• Best practice standards	11	12	
• Total other standards	32	31	
Total standards **Total numbered standards**	93 72	103 81	The total number of standards exceeds the total numbered standards because some standards have one or more best practice options and one has a transitional good practice option.

Methodology

9. This assessment assigned ratings for the MDBs' evaluation mandates, policies, guidelines and practices based on the information I had available as of February 28, 2010. It assigned two ratings for each standard. The first reflects the adoption of mandates, policies and guidelines materially consistent with the GPS. The second calls not only for the adoption of appropriate mandates, policies and guidelines but also for the substantially full application of the standards. The use of two ratings in this benchmarking report permits recognition of the extraordinary progress made by the IADB Group in revising its policies and guidelines during the course of the review and reflects the fact that these changes could not be fully integrated into its practices before the February 28, 2010, deadline.

10. This assessment applies the same rating system as in the Second Benchmarking review:[12]

[12] This approach can be criticized on at least two grounds, both valid. First, all standards receive equal weight even though some are clearly more important than others and even though a Not Materially Consistent (NC) rating on some standards requires NC ratings for others. For example, GPS 4.1.1 calls for assessing the MDB's additionality as a separate dimension of performance, distinct from the MDB's role and contribution. An MDB that does not meet the requirements of GPS 4.1.1 in this respect, cannot meet the requirements of GPS 4.5.1 and 4.5.2, which specify the indicators and benchmarks for rating additionality. Second, the results give an unjustified impression of precision. Alternative approaches, however, would also be subject to criticism. Seeking to agree weights to be assigned to each standard would have added a major, unnecessary complication and would have diverted the WGPSE's attention from more important matters. Provision of overall judgments not backed by systematic analysis could have been criticized as too subjective. Although the quantitative indicators in Chapter 2 should be seen only as broadly indicative of the extent of harmonization, rather than as precise measures, the approach used provides a useful summary of the extent of harmonization achieved.

Table 2: Rating System

Rating	Description
Materially Consistent (MC)	MDB's mandates, policies, guidelines and, where relevant, practices are materially consistent with the substance of the GPS.
Not Consistent (NC)	MDB's mandates, policies, guidelines and, where relevant, practices are not materially consistent with the substance of the GPS.
Not Relevant (NR)	(a) The MDB cannot follow the standard because of (i) the nature of its operations (not its policies or practices) *or* (ii) its total reliance on evaluations carried out by the central evaluation department (rather than evaluations by the MDB's operational staff), *or* (b) the MDB meets an equivalent or relevant higher standard specified in the GPS.

11. As in the prior assessments, I have calculated the consistency percentages excluding the standards not relevant for individual organizations.[13]

12. Throughout the benchmarking process, each member's evaluation department had extensive opportunities to comment on the draft harmonization matrix covering its institution. Based on a review of relevant documents and extensive communications with members during June-November 2009, I completed draft matrices, containing tentative ratings and the basis for these ratings, for each institution. In the course of this work, I provided extensive advice to the IADB Group with a view to improving the Group's private sector evaluation policies, guidelines and practices and helping them harmonize these policies, guidelines and practices with the GPS. In December 2009, the WGPSE decided that I should extend comparable advice to other institutions wishing to improve their policies and practices. Making this advice available to all reflected a recognition that the underlying objective is to improve and harmonize the member institutions' evaluation practices and that the benchmarking review is only a tool for encouraging this development. Consequently, the WGPSE decided to allow members until January 31, 2010, to submit to me for review proposals for changes and to put agreed changes into effect by February 28, 2010. I have taken these changes into consideration in assigning the final ratings. Since some members disagreed with some ratings, all members were given an opportunity to provide footnotes for insertion in this report.

2. PROGRESS ON GOOD PRACTICE HARMONIZATION STANDARDS

Overall Findings

13. Nearly all the surveyed member institutions have made significant progress. On average, they have adopted 78% of the good practice harmonization standards, and their practices were materially consistent with 69% of these standards, well above the 59% for the good practice harmonization standards in 2004 and 39% for the good practice standards in 2002.

[13] MIGA's evaluation unit has requested that the following footnote be included in the report:

> The Report's "consistency percentages" reflect some standards that in MIGA's view are not applicable to MIGA as an insurer. MIGA disagrees in several cases on whether a given standard for "investment operations" takes into account the differing circumstances and nature of MIGA's insurance business, and thus whether it is relevant for that organization . MIGA has noted via footnoting whenever the report's ratings are disputed on the basis of their applicability to the insurance business.

See Table 2 for definition of "not relevant," as agreed by the WGPSE. This definition distinguishes between the nature of an MDB's operations (e.g., an insurer that does not have authority to make equity investments cannot be expected to assign ratings for equity investments) and an MDB's policies and practices (e.g., its not requiring its clients to provide financial statements and, hence, its being unable to meet some standards).

14. The 2010 results, however, are not entirely comparable with the 2004 findings. As indicated in para. 5, the Third Edition introduced major changes in the standards. To permit comparisons, I looked separately at the ratings for the substantially unchanged standards and excluded MIF, which was not covered in 2004. The overall consistency ratings for adoption and substantially full application of the substantially unchanged standards increased from 64% to 81% between 2004 and 2010.[14]

15. The improvements reflect, among other things, the results of the second and third benchmarking reviews. In the aftermath of the 2004 review, which was distributed to some but not all boards of directors, IADB-SCF and MIF decided to embrace the good practice standards, and some other members made changes that improved their alignment with the standards. Subsequently, the discussions that took place during the current review led to further revisions in mandates, policies and guidelines to bring them into line with the standards. These changes, of course, must now be embodied in improved practices, and the next benchmarking review will show whether these changes are actually carried out.[15]

Findings, by Member

16. The dispersion of the scores has decreased greatly but still remains large. In 2004, the scores ranged from 8% to 92%. Now, they range from 46% to 98% for adoption of the standards and 44% to 93% for adoption and substantially full application of the standards. As was the case in 2004, the organizations operating solely or primarily in the private sector generally earned higher ratings than the two operating primarily in the public sector (Table 3 and Annex 1).

Table 3: Percentage of Materially Consistent Ratings for Harmonization Standards, By Member

Member	Overall 2010 Scores		Substantially Unchanged Standards		
			2010		2004
	Adoption	Adoption & Substantially Full Application	Adoption	Adoption & Substantially Full Application	Adoption & Substantially Full Application
AfDB	46%	44%	68%	65%	39%
AsDB	65%	63%	62%	59%	56%
EBRD	83%	81%	92%	89%	86%
IADB-SCF	90%	75%	94%	82%	8%
IIC	98%	79%	100%	91%	86%
MIF	75%	48%	78%	59%	n.a.
IFC	93%	93%	100%	100%	100%
MIGA	73%[16]	73%[17]	82%	82%	81%

[14] Comparability is limited to a lesser extent because I assessed performance more rigorously than in previous reviews. MIGA's evaluation unit has requested that the following footnote be included in the report with respect to this footnote:

> In MIGA's view, the more rigorous assessment of performance led to standards being applied inflexibly and in a way that did not take into account the differing circumstances and nature of MIGA's insurance business. MIGA footnotes indicate whenever the report's ratings are disputed on the basis of their applicability to the insurance business.

[15] Some members, however, may wish to wait for issuance of the Fourth Edition before changing their practices.

[16] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA has questioned several ratings in Annex 1 from the perspective of the relevance of those particular designed for IFI's making loans and equity investments to an insurer of political risk On the basis of MIGA's Comments on GPS standards #4.1.1, #4.3.2, #4.3.4, #4.5.1, and #4.5.2 (see MIGA's footnotes with MIGA Comments in Annex 1), MIGA believes the number of MC ratings (adoption and correct application) should be 42, and that **the percentage of MC ratings (adoption and correct application) should be 82% (not 73%).**

Overall	78%	69%	85%	81%	64%

N.B. Excludes experimental standards and excludes MIF from overall totals for substantially unchanged standards.

17. AfDB received scores of 46% for adoption of the standards and 44% for adoption and substantially correct application of the standards. Although it improved its consistency rating for the substantially unchanged standards from 39% in 2002 to 65% in 2010 (on an adoption and application basis), it has not yet adopted any of the new or substantially changed standards. AfDB may not have devoted sufficient attention to the private sector GPS, since private sector operations accounted for only 26% of its 2008 investments. It has expressed its intention to improve its performance.

18. AsDB scored 65% for adoption of the standards and 63% for adoption and substantially full application of the standards. AsDB's shortfalls included eight standards relating to annual reporting, seven where it had not taken the steps needed to adopt the GPS, and two where its concern with consistency with the rest of AsDB kept it from adopting the GPS standards. Private sector operations accounted for only 18% of AsDB's 2008 investments. AsDB has indicated its intention to improve its performance but has noted that its ability to do so is limited because sovereign lending accounts for the bulk of its operations.

19. EBRD received an 83% score for adoption of the standards and an 81% score for adoption and substantially full application. It disagreed with seven standards (three unchanged and four changed in the Third Edition). In addition, as indicated in para. 4, this benchmarking review has assigned not relevant ratings for six standards based on EBRD's statement that these standards, relating to business success and economic development, were not applicable to its operations.

20. The IADB Group's SCF did not begin to adopt the good practice standards until after the 2004 benchmarking review but, because of the work it did during 2005-09 and changes undertaken in the course of the current review, SCF received scores of 90% for adoption of the standards and 75% for adoption and substantially full application of the standards. Aside from not having had time to implement the revised guidelines, SCF did not meet the requirements of two standards relating to the evaluation mandate from IADB's Board and two standards on reporting to IADB's Board.

21. IIC received a 98% score for adoption of the standards and a 79% score for adoption and substantially full application of the standards. It made major improvements in its guidelines in the context of the benchmarking review and, though it did not have time to incorporate the changes into its practices, is seeking to apply them fully in the coming year.

22. Similarly, because of the work it did during 2005-09 and changes undertaken in the course of the current review, the IADB Group's MIF received scores of 75% for adoption of the standards and 48% for adoption and substantially full application of the standards. MIF's shortfalls arose mainly from its not having had time to implement the changes, not having a mandate approved by its governing body, and weaknesses in identifying the population and selecting the sample of operations to be evaluated.

23. IFC received a 93% score in both columns. It has adopted all of the unchanged standards but did not adopt four revised standards because it disagreed with them. Three involved IFC's rating additionality as part of work quality, rather than as a separate dimension of performance, and one involved the methodology for assessing the business success of loans to financial intermediaries.

---

[17] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA has questioned several ratings in Annex 1 from the perspective of the relevance of those particular designed for IFI's making loans and equity investments to an insurer of political risk On the basis of MIGA's Comments on GPS standards #4.1.1, #4.3.2, #4.3.4, #4.5.1, and #4.5.2 (see MIGA's footnotes with MIGA Comments in Annex 1), MIGA believes the number of MC ratings (adoption and correct application) should be 42, and that **the percentage of MC ratings (adoption and correct application) should be 82% (not 73%).**

24. MIGA received a score of 73% in both columns.[18] Among other things, it disagreed with the four standards that IFC rejected plus one other, and it fell short of the requirements for four standards relating to its annual review of evaluation results.[19]

Findings, by GPS Section

25. The average consistency ratings by GPS section range from 61% to 95% on an adoption basis or 61% to 85% on an adoption and substantially full application basis (Table 4 and Annex 1).

Table 4: Percentage of Materially Consistent Ratings for Harmonization Standards, By GPS Section

GPS Section (& % of Harmonization Standards)		Overall 2010 Scores		Substantially Unchanged Standards (Excluding MIF)		
				2010		2004
		Adoption	Adoption & Substantially Full Application	Adoption	Adoption & Substantially Full Application	Adoption & Substantially Full Application
1.	Roles of independent & self evaluation (11%)	77%	77%	91%	91%	71%
2.	Evaluation timing, population, coverage & sampling (13%)	77%	73%	84%	79%	60%
3.	Guidelines, execution and validation (9%)	95%	85%	94%	86%	55%
4.	Evaluative scope (54%)	79%	66%	88%	84%	61%
5.	Annual reporting & process transparency (13%)	61%	61%	62%	62%	65%
6.	Identification of lessons, dissemination, and ensuring application of lessons (0%)	No harmonization standards.				
Overall (100%)		78%	69%	85%	81%	64%

N.B. Excludes experimental standards.

26. *Roles of independent and self-evaluation—ensuring the evaluation department's independence.* The overall consistency rating for this section was 77%. Of the eleven not consistent ratings, five resulted from MIF's not having a mandate approved by its governing body, three from members' mandates not specifying that their Boards have ultimate authority over matters such as dismissing the head of the central evaluation department and deciding on his or her pay increases, and two from members' mandates not providing for the

[18] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA has questioned several ratings in Annex 1 from the perspective of the relevance of those particular designed for IFI's making loans and equity investments to an insurer of political risk On the basis of MIGA's Comments on GPS standards #4.1.1, #4.3.2, #4.3.4, #4.5.1, and #4.5.2 (see MIGA's footnotes with MIGA Comments in Annex 1), MIGA believes the <u>number</u> of MC ratings (adoption and correct application) should be 42, and that **the <u>percentage</u> of MC ratings (adoption and correct application) should be <u>82%</u> (not 73%).**

[19] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA disagrees with this assessment. MIGA notes that this benchmarking review assigned ratings for four standards which IEG-MIGA believes are not applicable to its insurance operations. IEG-MIGA believes that GPS standards #4.3.3, #4.3.5, #4.4.1, and #4.4.2 should be rated **<u>NR, and has conveyed this to the consultant.</u>**

evaluation department to have unrestricted access to clients and project sites. The shortfalls were limited to the IADB Group and AfDB.

27. *Evaluation timing, population coverage and sampling—ensuring that each institution evaluates a sufficiently large random sample drawn from the entire population of projects reaching early operating maturity.* The overall score for adoption and substantially full application of the standards under this section was 73%. Three institutions fell short of the requirements of the standard relating to defining the population, either because of not incorporating material elements of the GPS or because of mistakes in applying the GPS. Two did not adopt the GPS that defined early operating maturity for financial markets projects (including one that disagreed with the standard). Four did not meet the requirements of the standard relating to sample size (including one that disagreed with the standard).[20] Two did not meet the requirement calling for the central evaluation department, rather than the operating departments, to select the sample. And three did not provide information allowing a judgment on whether the samples were randomly selected.

28. *Guidelines, execution and validation—ensuring the quality of the reports that provide the base for the evaluation system.* This section, which showed the greatest improvement for the substantially unchanged harmonization standards, also showed the highest average ratings for adoption and for adoption and substantially full application of the standards (95% and 85%). The not consistent ratings arose mainly from shortfalls in the independent reviews of self-evaluation reports by the central evaluation departments of the IADB Group and AsDB and shortfalls in requiring and carrying out field visits.

29. *Evaluative scope—ensuring that the members evaluate the same performance dimensions, use the same performance indicators and adopt appropriate benchmarks.* This section—which accounts for over half the individual harmonization standards and embodies the essence of the harmonization effort—showed overall scores of 79% for adoption of the standards and 66% for adoption and substantially full application of the standards. The adoption and application score takes into account 74 not consistent ratings, spread over 22 of the 29 standards in this section. The substantially changed standards, representing 31% of the total number of standards in this section, accounted for 70% of the not consistent ratings for adoption and application. None of the substantially changed standards was adopted and applied by all the organizations covered by this review. In contrast, 44% of the substantially unchanged standards were adopted and applied by all.

30. The not consistent ratings resulted partly from substantive disagreements on seven standards. IFC and MIGA, for example, disagreed with the WGPSE's 2006 decision to rate additionality as an independent evaluation dimension, instead of including it as one of the components of work quality.[21] They also

---

[20] MIGA's evaluation unit has requested that the following footnote be included in the report:

> The shortfall in coverage is inherent in IEG-MIGA resourcing. While IEG-MIGA draws the GPS-required 50% random sample of guarantees annually for evaluation, the pool of evaluated guarantees in aggregate is not sufficient to comply with the good practice standards which require it to be statistically representative of the portfolio, at the 95% confidence level. (Compliance would imply a near 100% coverage of the population of MIGA projects underwritten each year, for several years, ie, until the pool of evaluations is large enough for statistical representativeness). IEG-MIGA is not resourced to do this, and it is unlikely that it will be in future. MIGA will revisit this issue as the GPS4 are developed

[21] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA was not part of the WGPSE's 2006 decision-making process and thus had no voice in the decision that additionality should be rated as an independent evaluation dimension instead of being included as a component of work quality. While MIGA does not have an indicator actually called "*Additionality",* in substance this aspect is rated and reported as *"MIGA's Role and Contribution*". MIGA believes that "*MIGA's Role and Contribution,*" which MIGA does rate separately, is substantively the same as "MIGA's additionality".

With respect to MIGA's statement that it "was not part of the WGPSE's decision-making process," see my comment on footnote 3. With respect to MIGA's belief "that *'MIGA's Role and Contribution*'...is substantially the same as 'MIGA's additionality,'" the WGPSE specifically decided at its March 2005, meeting that the draft of the Third Edition should, among other things, deal with splitting "additionality from work quality," and in its March 2006 meeting it approved the Third Edition, including this split. The WGPSE considered that "additionality" addressed an issue distinct from work disagreed with the practicality of WGPSE's 2006 decision to rate the business success of credit line operations directly, based on the project portfolio's contribution to the intermediary's financial returns, rather than indirectly, by looking at sub-project profitability and credit quality. Also, EBRD disagreed with assessing contributions to its own profitability based on benchmarks set in relation to overall corporate standards for minimally satisfactory expected performance, rather than using the *ex-ante* projections for each investment as the benchmark. EBRD's Board does not accept the use of corporate-wide hurdle rates.

31. Notwithstanding the substantial progress made, some of the "not consistent" ratings raise questions about some members' commitment to harmonization. For example, as indicated earlier, AfDB has not yet adopted any of the new or changed standards in the GPS, Third Edition. Similarly, EBRD has retained its earlier indicators or benchmarks with respect to environmental performance, additionality, and role and contribution, rather than adopting the benchmarks introduced in the GPS, Third Edition. And AsDB has considered that it had to maintain internal comparability on environmental and social ratings between its public and private sector evaluations, rather than harmonizing its standards with the other MDBs.

32. The developments at the IADB Group make the importance of commitment particularly clear. IADB's SCF and MIF had taken no steps to adopt the GPS before the second benchmarking review . Following that review, the IADB Group decided to adopt the standards but, in doing so, it fell short in many respects, partly due to efforts clarify or to simplify the standards. In the context of the discussions during the third benchmarking review, the IADB Group expressed its desire to bring its guidelines into line with the GPS. It accepted my offer to help it do so, and the revisions it made brought its scores for adoption of the standards in this section to 100% for IIC, 96% for IADB-SCF and 92% for MIF. Because it was too late to incorporate these changes in evaluation work already completed, the scores for adoption and substantially full application were lower, but I believe that these scores are likely to improve by the time of the next benchmarking review.

33. The WGPSE extended the timing for this review to allow other members an opportunity to revise their guidelines, but few members took advantage of this opportunity. Some central evaluation departments may have considered the short-term costs of changing their own approaches to be greater than the broader benefits of adopting the GPS. Changing their own guidelines would entail the cost of convincing others within the organization to adopt the new standards, the cost of reduced comparability between private sector and public sector evaluations within the institution, and the cost of revising earlier evaluation ratings to provide internal comparability over time. The benefits, in their view, may have been less clear.

34. *Annual reporting and process transparency—ensuring adequate reporting on results and on the evaluation process.* The overall consistency rating for this section was 61%, the lowest of the five sections with harmonization standards. Moreover, this section was the only one for which the average score for the substantially unchanged standards fell (albeit just slightly) from the level in the second benchmarking review.

35. These overall scores, however, mask wide disparities in performance. Three members met all the standards in this section; two met none. The not consistent ratings resulted from AfDB's not having issued annual reviews and annual reports, MIGA's recent annual reports' (in contrast to its earlier ones) not having

---

quality. Work quality is rated based on at-entry screening, appraisal and structuring; monitoring and supervision; and role and contribution. Role and contribution, in turn, is rated based on "compliance with basic operating principles, the MDB's contribution to client capacity building objectives (as relevant), the operation's consistency with furtherance of the MDB's corporate, country and sector strategies, and its clients' satisfaction with the MDB's service quality." Not only did the WGPSE recognize that the concept of "additionality" differed significantly from the concept of "work quality," but also it recogized that the MDBs' boards of directors had a high level of interest in additionality and deserved an independent, *ex post*, assessment of the additionality of operations evaluated, separate from other considerations.

covered all evaluations carried out during the year,[22] and shortfalls in the scope of coverage by AsDB, SCF, MIF and MIGA.[23] No issues of principle are involved here, and the members should all be able to follow these standards.

36. *Identification of lessons, disseminating, and ensuring application of lessons.* This section of the GPS includes no harmonization standards.

3. EXPERIMENTAL STANDARDS

37. The Third Edition introduced ten experimental standards. Seven specified benchmarks for rating contributions to business success, economic development and private sector development (GPS 4.3.3-6, 4.3.9, and 4.3.14-15). Two laid out innovative approaches for cost-benefit assessments where normal financial and economic returns cannot be calculated (GPS 4.3.7 and 4.3.13). And one sought to deal with the weaknesses in an earlier approach for assessing the contributions to economic development of credit lines and investments in equity funds (GPS 4.3.12). All aimed at closing gaps that limited the comparability of assessments of operational results sought by the Development Committee's 1996 Mandate.

38. Because of the innovative elements in these standards and the lack of consensus on them, the Working Group decided in 2006 to treat them as experimental. The members were to "experiment with these standards and then consider whether to retain them and, if so, whether to revise them...."[24] This chapter reports on the extent to which the members have adopted these standards.

39. In view of the innovative elements and the initial lack of consensus, it is not surprising that the experimental standards were adopted to only a modest extent. On average, the members adopted 46% of the standards, but the range went from zero to 100%. Three organizations adopted all ten of the experimental standards but did so too recently to permit an assessment of their application of the standards. Two others adopted some of the standards but rejected others because of substantive disagreements. The three remaining members received no materially consistent ratings for the experimental standards because they disagreed with the standards, did not incorporate all essential elements, or else changed the wording in ways that may

[22] MIGA's evaluation unit has requested that the following footnote be included in the report:

> MIGA holds that its recent Annual Reports ensured adequate reporting on results and on the evaluation process:
>
> (a) Specifically at the request of CODE, IEG-MIGA's 2008 Annual Report, *Evaluating MIGA's FY05-08 Strategic Directions*, was designed to inform the agency's new operational directions for FY09-11. It recommended a strategic framework that would take into consideration MIGA's operating environment, and the links between financial results, development impact and MIGA's strategic objectives. It highlighted the need for key performance indicators, for innovation to ensure that MIGA's product remains relevant, and for progress on development impact assessment and self-evaluation of projects. (b) IEG-MIGA's *2007 Annual Report* examined MIGA's systems for quality assurance, monitoring, and self-evaluation, and assessed the quality of the work in underwriting guarantee projects to ensure development effectiveness in MIGA's operations. The report's findings were based on quality at entry assessments of 50 MIGA projects underwritten between January 2005 and June 2006, using MIGA's policies and guidelines as benchmarks.

MIGA's annual reports may have been useful, but they did not meet the requirements in the GPS.

[23] MIGA's evaluation unit has requested that the following footnote be included in the report:

> The shortfall in coverage is inherent in IEG-MIGA resourcing. While IEG-MIGA draws the GPS-required 50% random sample of guarantees annually for evaluation, the pool of evaluated guarantees in aggregate is not sufficient to comply with the good practice standards which require it to be statistically representative of the portfolio, at the 95% confidence level. (Compliance would imply a near 100% coverage of the population of MIGA projects underwritten each year, for several years, ie, until the pool of evaluations is large enough for statistical representativeness). IEG-MIGA is not resourced to do this, and it is unlikely that it will be in future. MIGA will revisit this issue as the GPS4 are developed.

The GPS Section on reporting calls for the annual review to cover all the self-evaluation reports generated and reviewed by the central evaluation unit during the year. It does not concern itself with the number of projects evaluated.

[24] "Good Practice Standards for Evaluation of Private Sector Investment Operations, Third Edition, *op. cit*., p. 2.

have sought to improve on the experimental standards but made their guidelines inconsistent with them. Contrary to the 2006 expectations, few members actually experimented with any of the standards.

40. Clearly, many of the experimental standards will need revision in the next edition of the GPS. In revising these standards, however, the Working Group should keep in mind that comparability of rating judgments implies the need for common evaluation methods and benchmarks. If rating judgments are to be comparable, "satisfactory" must mean the same thing for each of the members. Equally important, the Working Group should keep in mind the importance of adopting evaluation methods and benchmarks relevant to the specific parameter being assessed. For example, sub-loan *financial* performance may provide some useful information but cannot, taken alone, support judgments on contributions to *economic* development. In revising the experimental standards, the Working Group should ensure that it does not issue standards that lead to misleading ratings.

41. I prepared a separate paper on the adoption of the experimental standards, the members' views and recommendations, and my conclusions.[25] Annex 2 to the present paper includes specific recommendations on these and the other (non-experimental) standards.

4. COMPARABILITY OF RATING CATEGORIES

42. An alternative way of looking at harmonization is to consider the comparability of the rating categories used by the members. The question here is whether the members are speaking the same language--whether, for example, "development outcome" or "additionality" mean the same things in different institutions.

43. The results--fourteen years after the owners called for harmonization of performance indicators--are disappointing.[26] Most of the members use the GPS terminology to describe their ratings categories, but (as shown in Annex 3) the content of these categories differs more often than not.

44. All the members assess development or transition outcome, but four define these categories differently than the GPS. When AfDB rates development outcomes, it does not take into consideration business success, one of the four indicators for rating development outcome in the GPS. (AfDB rates business success separately.) When AfDB and AsDB rate environmental and social performance, one of the four indicators for development outcome, they do not take into consideration the broader issue of the company's impacts in the vicinity of the project, as required by the GPS. When EBRD rates "transition outcome" (intended to be comparable to the GPS' development outcome), it takes into consideration "transition impact" (broadly equivalent to the GPS' contributions to private sector development), environmental performance and change (broadly equivalent to the GPS' environmental performance but also including environmental change), project and company financial performance (some overlap with the project's contribution to the company's business success in the GPS), and fulfilment of objectives (a secondary component in the GPS's assessment of contributions to the company's business success). As indicated in para. 4, however, EBRD has indicated that it seeks to foster transition, rather than development and poverty reduction, and that the standards relating to business success and economic development are, thus, not relevant for its operations. Since its coverage of components of development outcome diverges from the GPS, its transition outcome rating is not entirely consistent with the GPS's development outcome rating. And when MIGA rates contributions to the

[25] Walter I. Cohn & Associates, LLC, "Experimental Standards in MDB-ECG Good-Practice Standards for Evaluation of Private Sector Investment Operations, 3d Ed.," April 12. 2010.

[26] MIGA's evaluation unit has requested that the following footnote be included in the report:

> IEG-MIGA notes that it has been in existence only since July 2002 (7 years), and believes that, far from "disappointing results", much has been achieved in this period.

company's business success, it considers a project's contributions to other business goals articulated at approval only when no quantitative information is available and does not consider the project company's prospects for sustainability and growth, as specified by the GPS.

45. Although nearly all the members assess contributions to the MDB's own profitability, two use definitions that differ from the GPS. AsDB allows the operating staff to choose among three proxies for "at-approval standards," and clearly all cannot be equivalent to "at-approval standards for minimally satisfactory expected performance." EBRD bases its ratings on the projections for the individual projects at approval, rather than general, corporate standards for minimally satisfactory expected performance.

46. Five members call for rating additionality as a separate dimension, but three (AsDB, EBRD and MIF) use benchmarks that limit comparability with the GPS. Three others (AfDB, IFC and MIGA) continue to include additionality as a component of work quality, rather than as a separate dimension.

47. All the members assess work quality, but five use definitions that differ from the GPS. AfDB, IFC and MIGA continue to include additionality as a component of work quality, and AsDB and EBRD use definitions of the MDB's role and definition that differ from the GPS.

5. CONCLUSIONS AND RECOMMENDATIONS

48. Nearly all the organizations covered have made significant progress towards harmonization. The second benchmarking review--which had been circulated to the boards of directors of most members--stimulated some to embrace the good practice standards, bring more of their practices into line with the standards, or both. In addition, the advice provided during the current benchmarking review helped several organizations to correct shortfalls in their evaluation mandates or guidelines. The recent changes, of course, still need to be incorporated in the members' practices.

49. Notwithstanding this progress, the members still have a way to go. Fourteen years after the owners called for harmonization, the members' scores should all by now be approaching 100%. Most of the members fall short of this objective, including a few who fall well short of it. Moreover, although most of the members use the GPS terminology to describe their ratings categories, the content of these categories often differs. The members may use the same words for their ratings dimensions and indicators, but the meaning of the words often differs from one institution to another. And despite the progress on benchmarks and methodological issues, the members have not yet achieved consensus on the experimental standards, which are essential to permit comparable assessments of contributions to business success and economic development, two of the four indicators for development outcomes.

50. Two factors account for most of the shortfalls in harmonization. First, some members disagree with some standards. This problem applies particularly to Section 4 of the GPS, which deals with the evaluative dimensions, indicators, and benchmarks. It is, of course, especially acute for the experimental standards, which are new and innovative. Some of these disagreements might have been avoided had the Working Group allocated sufficient time for resolving them before issuing the GPS.

51. Second, some members have not adequately committed themselves to the objective of harmonization. Four years after issuance of the Third Edition, several members still retain the second edition standards. Several did not take advantage of the opportunity to make changes after I provided advice on what would be needed to bring their mandates, guidelines, and practices into line with the standards. Some paid inadequate attention to the standards in defining the population for possible evaluation and selecting the sample of projects for evaluation. And some ignored important elements of the standards on annual reporting.

52. Change requires time and effort. Some evaluation departments may have been unwilling to undertake the efforts needed to convince others within their institutions (possibly, in some cases, including their boards of directors) to adopt new standards, to insist on the need for using different standards for evaluating public and private sector operations, or to retrofit earlier evaluation ratings to permit internal comparability over time.

53. ***The members need to pursue their efforts towards harmonization, not only to fulfill the mandate laid down by the owners but also--and equally important--to ensure that their evaluation work is consistent with practices that will yield the most reliable, most useful results. To this end,***

- ***The next edition of the GPS should build on the progress already made and further clarify and strengthen the standards. It should not accept practices that can lead to misleading ratings on development outcomes, the MDB's investment profitability, the MDB's additionality, or the MDB's work quality; misleading judgments on overall results; and unjustified conclusions for future operations. Annex 2 provides suggestions for possible changes in the standards.***
- ***The WGPSE should allocate sufficient time to reach consensus on the next edition of the GPS, including revisions to the experimental standards. Before the WGPSE meets to discuss the draft Fourth Edition, it should provide a period of several months for written comments and exchanges of views. Expecting to reach consensus in a single meeting, lasting less than a day, has proven to be unrealistic.***
- ***To facilitate consensus on performance benchmarks, the good practice standards should be limited to the benchmarks for "satisfactory" ratings and should treat the other benchmarks as best practice standards.***
- ***The members who are lagging should explicitly commit themselves to bringing their mandates, guidelines and practices into line with the GPS by the next benchmarking review. They may be subject to some constraints that appear insuperable but should seriously consider what might be done to overcome these constraints.***
- ***The next benchmarking review should allow sufficient time to assess whether a reasonably sized random sample of each member's evaluation reports is consistent with the GPS, i.e., whether the members are correctly applying the standards in practice. Correct application of the standards will lead not only to greater harmonization but--equally important--to evaluations that are more solidly grounded and, hence, more useful.***[27]

August 25, 2010

[27] MIGA's evaluation unit has requested that the following footnote be included in the report:

> With the development of GPS4, these standards are being reconsidered by the WGPSE and full application of GPS3 standards may soon no longer be the goal. The changes suggested in this paragraph would be premature, pending the revision of the standards that is underway.

It would be a pity if the member institutions were to decide to weaken their efforts to strive for harmonization of their evaluation practices. They have made great progress towards harmonization and, in doing so, have significantly improved their practices and the value of their evaluation products.

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL APPLICATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
1	**Roles of Independent and Self-Evaluation** (Out of 47)	**3**	**6**	**6**	**4**	**5**	**1**	**6**	**5**	**36**	**77%**
1.1	***Structure and role of independent evaluation, i.e., of Central Evaluation Department (CED)*** *(Out of 5x8=40)*	3	5	5	3	4	0	5	5	30	75%
1.1.1	CED has a Board-approved mandate statement, designed to ensure independence and relevance.	MC	MC	MC	MC	MC	NC	MC	MC	7	7/8
1.1.2	The mandate provides that the Board of Directors oversees CED's work and that the CED's reporting line, staff, budget and functions are organizationally independent from the MDB's operational, policy and strategy departments and related decision-making. In addition, the mandate specifies that the Board has the ultimate decision authority for (1) hiring and terminating CED head and staff; (2) CED head's appointment terms and reporting structure; (3) CED head's and staff's grading, performance reviews and pay increases; and (4) the CED's budget.	NC+	MC	MC	NC+	NC+	NC +	MC	MC	4	4/8
1.1.4	Under its mandate, CED's scope of responsibility extends, without restriction, to all determinants of the MDB's operational results.	MC	MC	MC	MC	MC	NC	MC	MC	7	7/8
1.1.5	The mandate states that CED has unrestricted access to MDB's staff, records, co-financiers, clients and projects. The mandate may, however, allow for restrictions on access to clients and projects in jeopardy cases, where an evaluator's visit could prejudice the MDB's financial interests or materially increase the risk of litigation. Should client access be restricted in jeopardy cases, the number of such cases should be reported in the MDB's annual report or annual review.	NC+	MC	MC	NC+	MC	NC	MC	MC	5	5/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
1.1.6	The mandate provides that CED transmits its reports to MDB's Board after review and comment by management but without management clearance or any management-imposed restrictions on their scope and contents.	MC	MC	MC	MC	MC	NC	MC	MC	7	7/8
1.2	***Responsibilities of operations departments in self-evaluation--****Out of 1x7=7*	0	1	1	1	1	1	1	0	6	86%
1.2.1	*Good practice.* Execute XASRs on investments selected pursuant to GPS 2.3.1 and 2.3.2 in accordance with CED's sample selection and evaluation guidelines. *Best practice.* In addition, deliver XASRs according to a schedule designed to spread the review load throughout the program year and allow CED to complete the annual review on schedule. 1.2.1	NC MC-	MC NC	MC MC	MC NC	MC MC	MC NC	MC MC	NR NC	6 4	6/7 4/8
2	**Evaluation Timing, Population, Coverage and Sampling** (Out of 52)	**2 (3)**	**4 (5)**	**6**	**5**	**5**	**3**	**7**	**6**	**38 (40)**	**73% (77%)**
2.1	***Identification of population from which sample for evaluation is to be drawn; timing of consideration for evaluation*** *(Out of 4x8=32)*	2	3	3	4	4	3	4	4	27	84%
2.1.1	Taking into consideration information on project maturity status provided by other departments, CED determines the *population* from which the investments to be evaluated each year are to be drawn.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
2.1.2	The population from which the investments to be evaluated each year are to be drawn consists of the investments that will have reached *early operating maturity* (as defined in GPS 2.1.3 and 2.1.4) during the year. • Subject to certain exclusions, specified below, the population includes all disbursed (including partially cancelled) investments--whether still active or already closed (paid-off, sold or written off)--that have reached early operating maturity. The population also includes investments already closed, even if they never reached early operating maturity. • Excluded from the population are dropped and cancelled investments, very small investments made under special promotional programs (e.g., direct investments in small businesses that are evaluated on a program basis through a CED special study), subscribed rights offerings and investments undertaken to help finance cost overruns on projects previously financed by the MDB. • Projects that have not yet reached early operating maturity are excluded from the current evaluation year's population and rolled forward for inclusion in the population in a future year when they will have reached early operating maturity. • Investments are included in the population from which the sample for evaluation is drawn only once, i.e., only for the year in which they will have reached early operating maturity.	NC	NC+	MC-	MC	MC	NC	MC-	MC	5	5/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
2.1.3	All operations other than the financial markets operations specifically covered by GPS 2.1.4 are deemed to have reached *early operating maturity* when (a) the project financed will have been substantially completed, (b) the project financed will have generated at least 18 months of operating revenues for the company and (c) the MDB will have received at least one set of audited annual financial statements covering at least 12 months of operating revenues generated by the project.	MC-	MC-	MC-	MC	MC	MC	MC	MC	8	8/8
2.1.4	Financial markets projects where the principle objective is to assist identifiable capital expenditure sub-projects (rather than to contribute to institutional development or institution building) are deemed to have reached *early operating maturity* after the elapse of at least 30 months following the MDB's final material disbursement for sub-loans or sub-investments, i.e., ignoring disbursements for small follow-up investments in existing client companies and disbursements to cover management fees or other expenses of investment funds.	NC	MC	NC	MC-	MC	MC	MC	MC	6	6/8
2.2	***Evaluation coverage*** *(Out of 1x8= 8)*	0	0 (1)	1	1	1	0	1	0	4 (5)	50% (63%)
2.2.1	***N.B. MC for any of these 4 alternatives is equivalent to MC for good practice standard.*** *Good practice:* Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample of sufficient size to establish, *for a combined three-year rolling sample*, success rates at the 95% confidence level, with sampling error not exceeding ±5 percentage points, for the population's development (transition) outcome, MDB investment outcome, additionality and MDB work quality.	NC	NC (MC)	MC	NC	NC	NC	MC	NC[1]	2 (3)	2/8 (3/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
	Transitional good practice: Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample equivalent to 60% or more of the investments in the population. In using this standard, an MDB reports on the confidence level and sampling error applicable to the success rates for the population's development or transition outcome, MDB investment outcome, additionality and MDB work quality. An MDB can use this standard only until its combined three-year rolling population of projects reaching early operating maturity reaches 50.	NC	NC	NC	NC	NC	NC	NC	NC	0	0/8
	Best practice-Alternative 1: Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample of sufficient size to establish, *for a combined three-year rolling sample*, success rates at the 95% confidence level, with sampling error not exceeding ±5 percentage points, for the population's development or transition outcome, MDB investment outcome, additionality and MDB work quality *within each of the MDB's current strategically targeted groups.*	NC	NC	NC	NC	NC	NC	NC	NC	0	0/8
	Best practice-Alternative 2: Preparation of XASRs (with XASR-As), PERs, or a combination of the two on 100% of the investments in the population.	NC	NC	NC	MC	MC	NC	NC	NC	2	2/8
2.3	***Sampling*** *(Out of 2x6 = 12)*	0 (1)	1	2	0	0	0	2	2	7 (8)	58% (67%)
2.3.1	The CED selects the operations for XASRs and PERs from the evaluation year's population (as defined above), subject to the following standard.	NC (MC)	MC	MC	NR	NR	NC	MC	MC	4 (5)	4/6 (5/6)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
2.3.2	If coverage is less than 100%, the sample should be both random and representative. Notwithstanding this principle, a CED may wish to select projects to be covered by PERs based, e.g., on the potential for learning, the high profile of an operation, credit and other risks, whether the sector is a new one for the MDB, the likelihood of replication, or the desirability of balanced country and sector coverage. If so and if the CED wishes to combine the PER with the XASR-A findings in reporting annual success rates, it uses stratified sampling methodology, as follows: • It splits the population into two strata. The first consists of the projects CED selects for PERs. The second consists of the remaining projects, i.e., the population other than the projects selected for the first stratum. • CED evaluates 100% of the first stratum. The success rates from the sample are, thus, identical to the success rates for this portion of the population. • CED selects a random sample from the second stratum. The operational staff prepares XASRs on the projects selected, and CED prepares XASR-As on these projects (or a random sample of them). The sample is sufficiently large to give reliable estimates of the success rates for that stratum. • Based on the weight of each stratum in the overall population, CED then calculates the weighted average success rates and sampling errors, following the normal procedures for stratified sampling.	NC	NC	MC	NR	NR	NC	MC	MC	3	3/6

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
	If the CED wishes to select projects to be covered by PERs as above *but does not wish to combine the PER with the XASR-A findings in reporting annual success rates*, it draws the sample to be covered by XASRs from the full population for the year, without previously eliminating the projects to be covered by PERs. To the extent that specific projects may be selected for XASR-As and PERs, CED would use the PER ratings, rather than the XASR-A ratings in reporting on success rates, since CED will have carried out a more rigorous review in these cases.										
3	**Guidelines, Execution and Validation** (Out of 40)	**5**	**3**	**5**	**3 (5)**	**4 (5)**	**4 (5)**	**5**	**5**	**34 (38)**	**85% (95%)**
3.1	***Guidelines & familiarisation*** *(Out of 2 x 8 = 16)*	2	2	2	2	2	2	2	2	16	100%
3.1.1	In consultation with operations departments, CED prepares, refines and disseminates guidelines for the preparation of XASRs and PERs in sufficient detail to promote consistency and objectivity in execution scope, analysis and ratings. As part of dissemination efforts, some CEDs may wish to conduct workshops to familiarize the XASR teams with requirements and supporting documentation for achieving good-practice execution.	MC-	MC	MC	MC	MC	MC	MC	MC-	8	8/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)
(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
3.1.2	*Good practice*: The guidelines include ratings guidelines with benchmarks and standard reporting templates that include the performance ratings matrix. *Best practice*: The guidelines also include related documentation, such as an overview of the XASR program, a description of efficacious execution process steps, good-practice examples of XASRs from previous years' samples, and a list of execution mistakes to avoid (informed by past XASRs).	MC NC	MC NC	MC MC	MC MC	MC MC	MC NC	MC MC	MC MC	8 5	8/8 5/8
3.2	***Execution*** *(Out of 1x8 = 8)*	1	0	1	0 (1)	1	1	1	1	6 (7)	75% (88%)
3.2.1	The research for XASRs and PERs draws from a file review; discussions with available staff involved with the operation since its inception; independent research (e.g. on market prospects); a field visit to obtain company managers' insights and to the project site to observe and assess outcomes; and discussions with parties who are knowledgeable about the country, company and project (e.g. MDB specialists, company employees and auditors, suppliers, customers, competitors, bankers, any relevant government officials, industry associations, and local NGOs).	MC	NC+	MC	NC (MC)	MC	MC	MC	MC-	6 (7)	6/8 (7/8)
3.3	***Review and independent validation*** *(Out of 2x8 = 16)*	2	1	2	1 (2)	1 (2)	1 (2)	2	2	12 (15)	75% (94%)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
3.3.3	CED conducts an independent review of each XASR to verify scope responsiveness, evident reliability of the analysis, impartiality and consistency in ratings judgments, and appropriateness and completeness of the identified lessons, and then, *for each randomly selected XASR to be used in the annual synthesis report on evaluation results*, prepares an XASR-A on the final-edition XASR that records its independent judgments on the report's quality in relation to the guidelines, assigned ratings and lessons.	MC	NC+	MC	NC (MC)	NC (MC)	NC (MC)	MC	MC-	4 (7)	4/8 (7/8)
3.3.5	Following preparation of each draft XASR-A, CED reviews with the XASR team and its manager the basis for its judgments where its ratings differ from those in the final edition XASR.	MC	MC	MC	MC	MC	MC	MC	MC-	8	8/8
4	**Evaluative Scope** (Out of 219)	**14**	**21**	**15 (16)**	**21 (27)**	**20 (29)**	**10 (22)**	**25**	**19**	**145 (173)**	**66% (79%)**
4.1	***Performance dimensions evaluated*** *(Out of 3x8 = 24)*	2	3	2	3	3	2 (3)	2	2	19 (20)	79% (83%)
4.1.1	*Good practice*: The scope of the XASR (and XASR-A) or PER includes, at a minimum, • The *development or transition outcome of the project and, to the extent provided in GPS 4.3.3-4.3.7 and GPS 4.3.17, the project company*, i.e., the "results on the ground" relative to the MDB's mission. • The *MDB investment's profitability* (contribution to its corporate profitability objective),[2] • The *MDB's additionality, and* • The *MDB's work quality* (also referred to as bank handling, operational effectiveness, or execution quality).[3]	NC	MC	NC+	MC	MC	NC (MC)	NC+	NC+[4]	3 (4)	3/8 4/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
4.1.2	The operation's performance under each of these dimensions is analyzed according to standard indicators, and the operation's performance for each indicator is rated according to criteria and benchmarks specified in the guidelines.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.1.3	The performance reflected in the relevant indicator ratings is synthesized into ratings for each of the four performance dimensions, specified above.	MC-	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.2	***Performance ratings—principles and benchmarks*** *(Out of 5x8 = 40)*	5	4	5	4	5	5	5	5	38	95%
4.2.1	Within the rating scales (e.g., ranging from *unsatisfactory* to *excellent*), there should be balance between positive and negative characterizations (i.e., if there are four ratings, two are less than good and two are good or better). The words used to describe these ratings should accurately reflect whether the judgments are less than good or else good or better.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.2.2	Each of the evaluated performance attributes is assigned a rating using a 4- to 6-point scale for each indicator ratings and a 4- to 6-point scale for each synthesis rating.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.2.3	Ratings of non-quantitative indicators require that relative qualitative judgments be made. The criteria should reflect the extent to which performance has been consistent with the MDB's policies, prescribed standards for corporate sustainability and recognized good-practice standards. The criteria for the judgments should be clearly specified in the guidelines for the preparation of XASRs and in the CED's annual review.	MC	NC+	MC	NC+	MC	MC	MC	MC	6	6/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.2.4	The synthesis ratings for the four dimensions (development or transition outcomes, profitability to the MDB, the MDB's additionality and the MDB's work quality) reflect summary qualitative performance judgments based on the underlying indicator ratings. They are not simple averages of the indicator ratings.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.2.5	Outcomes for each of the indicators are assessed on a "with v. without project" basis.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.3	***Indicators and benchmarks for the development or transition outcome (the project's contribution to (a) the company's business success;(b) the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy; and (c) economic development; and (d) the project's overall environmental performance)*** *(Out of (3x8) + (3x7) + (4x6) = 69)*	1	8	3	6 (10)	4 (10)	1 (5)	9	6	38 (52)	55% (75%)
4.3.1	Assessments of development or transition outcomes for each of the development or transition outcome indicators take into consideration the sustainability of the results.	MC	MC	MC	MC	MC	MC	MC	MC	8	8/8
4.3.2	The project's development or transition outcome is based partly on the *project's contribution to the company's business success,* measured primarily by the real after tax returns and secondarily by (i) the project's contribution to other business goals articulated at approval and (ii) the project *company's* prospects for sustainability and growth.	NC	MC	NR	NC (MC)	NC (MC)	NC (MC)	MC	NC+ [5]	2 (5)	2/7 (5/7)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.3	*In rating the business success of capital expenditure projects where the incremental costs and benefits can be quantified,* the project's after-tax financial rate of return in real terms (FRR) is compared with the company's weighted average cost of capital (WACC). Financial and economic theory holds that a firm must expect an after-tax financial rate of return on the funds it invests that is at least sufficient to induce investors to purchase and hold the firm's debt and equity. The investors' return requirements are reflected in the company's WACC, i.e., the weighted average after-tax cost to the company of the yields it must provide on its borrowings and the equity investors' minimally acceptable returns, all adjusted for inflation. More specifically, the WACC is the sum of (a)(i) the average after-tax cost of the company's debt multiplied by (ii) the company's debt as a percentage of its debt plus equity; plus (b)(i) the cost of the company's equity multiplied by (ii) the company's equity as a percentage of its debt plus equity; minus (c) the inflation rate. In estimating the WAAC, the cost of the company's equity is assumed to be the average nominal pre-tax cost of the company's debt plus a an equity premium. ~~In estimating the WAAC, the cost of the company's equity is assumed to be the average nominal pre-tax cost of the company's debt plus a 350 bp equity premium.~~ ***N.B. The harmonization standard does <u>not</u> take into consideration the level of the equity risk premium. This component of the standard, the experimental component, has been crossed out here.***	NC	MC	NR	NC (MC)	NC (MC)	NR	MC	NC+[6]	2 (4)	2/6 (4/6)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.4	The ratings benchmarks for the real after-tax FRR for projects covered by this standard are: • Excellent: FRR exceeds the average nominal cost of the company's borrowings by ~~700 bp or more.~~ • Satisfactory: FRR is equal to or greater than the WACC but less than the FRR required for an excellent rating. • Partly unsatisfactory. FRR is equal to or greater than what the WACC would be if the shareholders earned the same return as the lenders (i.e., if the equity premium were zero, rather than ~~350 bp over the nominal pre-tax cost of the company's debt~~) but is less than the FRR required for a satisfactory rating. • Unsatisfactory. FRR is lower than the FRR required for a partly satisfactory rating (i.e., the shareholders earned less than the lenders). The business success ratings also take into consideration, where appropriate, (i) the project's contribution to other business goals articulated at approval and (ii) the project *company's* overall prospects for sustainability and growth. ***N.B. The harmonization standard does <u>not</u> take into consideration the level of the premium for an excellent rating and the level of the equity risk premium. These components of the standard, the experimental components, have been crossed out here.***	NC	MC	NR	MC	MC	NR	MC-	NC+[7]	4	4/6

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.5	*In rating the business success of operations involving loans to intermediaries to finance identifiable capital expenditure sub-projects,* the project portfolio's contribution to the after-tax real return on the intermediary's equity is compared with the equity returns implied by the FRR benchmarks calculated using the methodology outlined in GPS 4.3.3). If cost accounting data are available, the information needed to estimate the incremental return on the intermediary's equity can be derived from the intermediary's cost accounting system. When cost accounting data are not available, which will normally be the case, the MDB's or the CED's staff can help the financial intermediary develop rough *ad hoc* estimates of the rate of return. The after-tax real return on equity is calculated from (i) the actual (or typical) spreads and other charges on sub-loans financed by the MDB's loan, (ii) the principal amounts to which these charges would apply each year, (iii) write-offs and expected write-offs (i.e., end of projection period loss provisions) on sub-loans financed by the MDB's loan, (iv) administrative expenses, based on cost accounting data, *ad hoc* estimates, or educated guesses, e.g., as to how the average administrative costs for the sub-loans financed might compare with the institution's overall average administrative expenses, (v) collateral benefits, based on the intermediary's accounting system, *ad hoc* estimates, or educated guesses, (vi) adjustments for inflation and for gains or losses, if any, resulting from exchange rate changes, (vii) income taxes, and (viii) the share of the intermediary's equity needed to back the sub-loans, based either on the intermediary's or the BIS' risk weights. (Exhibit A provides a simplified example of how the return can be calculated.) The assumptions underlying the estimates should be specified in the XASR or an attachment.	NC	MC-	NR	NC (MC)	NC (MC)	NR	NC	NC[8]	1 (3)	1/6 (3/6)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
	For projects covered by this standard, the ratings benchmarks for the real equity returns are derived from the benchmarks in GPS 4.3.4 as follows: • Excellent: (a) The real equity return implied by the FRR required for an excellent rating under GPS 4.3.4. ~~(This benchmark is equivalent to the sum of (i) the WACC (as calculated in accordance with GPS 4.3.4) plus a 250 bp premium (required for an excellent FRR) minus (ii) the average after-tax cost of the company's total debt multiplied by the company's debt as a percentage of its total assets plus (iii) the inflation rate, all divided by (b) the company's equity as a percentage of its total assets.)~~ • Satisfactory: (a) At least (i) the average nominal pre-tax cost of the company's debt plus (ii) ~~350 basis points~~ but (b) less than the benchmark for an excellent rating. • but (b) less than the benchmark for an excellent rating. • Partly Unsatisfactory: (a) At least the average nominal cost of the company's debt but (b) less than the benchmark for a satisfactory rating. • Unsatisfactory: Below the average nominal cost of the company's debt. The business success ratings also take into consideration, where appropriate, (i) the project's contribution to other business goals articulated at approval (e.g., reducing the financial intermediary's riskiness) and (ii) the *intermediary's* overall prospects for sustainability and growth. ***N.B. The harmonization standard does <u>not</u> take into consideration the level of the premium for an excellent rating and the level of the equity risk premium. These components of the standard, the experimental components, have been crossed out here.***										

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.8	The project's development or transition outcome is based partly on the *project's contributions (positive or negative) to the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy.* In assigning ratings for this standard, the following factors may be considered: • Competition: The project contributes to greater efficiency, quality, innovation or customer orientation of other suppliers through competitive pressures or contributes to restrictions on competition, e.g., by increased protective tariffs, cartels, etc. • Market expansion: Expansion of markets through the project entity's interactions with suppliers (backward linkages) and customers (forward linkages) and through contributions to the integration of economic activities with the national or international economy. • Private ownership and entrepreneurship: Significant increase or consolidation of private provision of goods and services and support for entrepreneurial initiative; or weakening of support for private ownership and entrepreneurship due to allocation by a financial institution of resources to purchases of government securities or loans to state-owned enterprises. • Frameworks for markets (institutions, laws and policies that promote market functioning and efficiency): Creation or strengthening (or weakening) of public and private institutions that support the efficiency of markets; improvements to (or weakening of) the functioning of regulatory entities and practices; contributions (positive or negative) to government policy formation and commitment, promoting competition,	NC+	MC-	MC-	MC	NC (MC)	NC (MC)	MC	MC	5 (7)	5/8 (7/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
	• predictability and transparency; contributions to laws that strengthen (or weaken) the private sector and an open economy. • Development of financial institutions and financial markets: Contributions (positive and negative) to the development of sustainable financial institutions and the financial markets in which they operate (including creation of new fund management companies to manage the MDB-supported investment fund, creation of new fund management companies by staff from the management company responsible for the MDB-sponsored fund, and creation of subsequent investment funds); improved financial strength in sector (e.g., by improving asset-liability management); pioneering listing on stock exchange or significant broadening of listed value; first-of-a-kind financial instrument in local market; greater resource mobilization; and improved allocation efficiency. • Transfer and dispersion of skills: Project contributes to significant upgrading of technical and managerial skills beyond the project entity; introduction of new technology or know-how, including financial know-how. • Demonstration effects (spread of new behaviors and activities): Demonstration of replicable products and processes that are new to the economy; new investments stimulated by the project; demonstration of ways of successfully restructuring companies and institutions; demonstration of new ways and instruments to finance private sector activity. • Standards for corporate governance and business conduct: Improvements in (or weakening of) standards, e.g., with respect to accounting standards, disclosure standards, risk management standards and the company's governance quality, reputation and business practices (including corruption) as a positive (or negative) corporate role model and quality investment asset. • Development of physical infrastructure used by other private parties.										

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.10	The project's development outcome is based partly on its *contributions to economic development.* Performance is assessed not only by the direct economic costs and benefits to the owners and financiers but also by the economic costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	NC+	MC	NR	MC	NC (MC)	NC (MC)	MC	MC	4 (6)	4/7 (6/7)
4.3.11	*For capital expenditure projects where the incremental costs and benefits can be separately quantified*, the economic development assessment is based mainly on the project's net quantifiable economic benefits and costs, as measured by the project's real economic rate of return (ERR), but taking into consideration also other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	NC	MC	NR	NC (MC)	NC (MC)	NR	MC	MC	3 (5)	3/6 (5/6)
4.3.16	The project's development or transition outcome is based partly on the company's overall environmental performance at the time of evaluation. The assessment is based primarily on the MDB's specified standards in effect at investment approval (though compliance with the standards specified at the time of the evaluation may be taken into consideration in assigning "outstanding" or "excellent" ratings). "Environment" includes the physical environment and, to the extent covered by the MDB's policies, also includes energy efficiency; the quality of the client's environmental management activities; social, cultural, health and safety impacts; and the extent of public consultation and participation.	NC+	NC+	MC-	MC	MC	NC (MC)	MC	MC-	5 (6)	5/8 (6/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.3.17	The environmental performance is rated based substantially on the following benchmarks: For non-financial markets projects: • Excellent: The company (a) meets (i) MDB's at approval requirements (including implementation of the environmental action program, if any) and (ii) MDB's at evaluation requirements; and (b) has either (i) gone beyond the expectations of the environmental action plan or (ii) materially improved its overall environmental performance (e.g., through addressing pre-existing environmental issues) or (iii) contributed to a material improvement in the environmental performance of local companies (e.g., by raising industry standards, acting as a good practice example, etc.) • Satisfactory: The company is in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any). • Partly unsatisfactory: (a) The company is not in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any) but is addressing deficiencies through on-going or planned actions; or (b) earlier non-compliance (subsequently corrected) resulted in environmental damage that has not been corrected. • Unsatisfactory: (a) The company is not in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any), and (b)(i) mitigation prospects are uncertain or unlikely; or (ii) earlier non-compliance (subsequently corrected) resulted in substantial and permanent environmental damage. No opinion possible: Where, after best efforts, the relevant information to establish material compliance (or lack thereof) cannot be obtained, a rating of no opinion possible may be assigned. This rating should be a last resort, after reasonable effort has been made to obtain the necessary information. A sponsor's failure to report should result in a partly unsatisfactory or unsatisfactory rating only if the sponsor has repeatedly refused to cooperate on this issue.	NC+	NC+	NC+	MC	MC	NR	MC	MC-	4	4/7

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
4.4	***Indicators and benchmarks for MDB's investment profitability*** *(Out of (2x8) + (1x6) = 22)*	3	1	1	2	3	1 (3)	3	0	14 (16)	64% (73%)
4.4.1	*Good practice.* MDB investment's profitability is based upon the investment's *gross profit contribution* (net of financing costs and loss provisions but before deducting administrative costs) with ratings benchmarks set in relation to corresponding at-approval standards for minimally satisfactory expected performance.	MC	NC+	NC	MC	MC	NC (MC)	MC	NC+ [9]	4 (5)	4/8 (5/8)
	Best practice #1. If reliable transaction cost data are readily available from management information system, MDB investment's profitability is also based on the investment's *net profit contribution* (the gross profit contribution less administrative costs), measured in risk-adjusted, discounted cash flow terms with ratings benchmarks set in relation to at-approval standards for minimally satisfactory expected performance.	NC	NC	NC	NC	NC	NC	NC	NC+ [10]	0	0/8
	Best practice #2. If reliable transaction cost data are readily available from management information system, MDB investment's profitability is also based on the investment's *net profit contribution (the gross profit contribution less administrative costs) in relation to the capital employed for the investment,* measured in risk-adjusted, discounted cash flow terms with ratings benchmarks set in relation to at-approval standards for minimally satisfactory expected performance.	NC	NC	NC	NC	NC	NC	NC	NC [11]	0	0/8

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.4.2	*Good practice.* Loan and guarantee performance benchmarks are set in relation to the MDB's expectations at approval.	MC	MC	MC	MC	MC	MC	MC	NC+ [12]	7	7/8
	Best-practice: Loan's net profit contribution (net of transaction and financing costs) is sufficient in relation to the MDB's return on capital employed target.	NC	NC	NC	NC	NC	NC	NC	NR [13]	0	0/8
4.4.3	*Good practice.* Equity investment rates of return are benchmarked (or discounted if an NPV measure is used) against standards for minimally satisfactory expected performance at approval.	MC	NC+	NC	NR	MC	NC (MC)	MC	NR	3 (4)	3/6 (4/6)
	Best practice 1. Equity investment rates of return are benchmarked (or discounted if an NPV measure is used) to reflect appropriate spreads over actual or notional loan yields for the same credit risk, in line with the policy-defined, at-entry approval standard.	NC	NC	NC	NR	NC	NC	MC	NR	1	1/6
	Best practice 2. Where the MDB's investment features both a loan and an equity investment, their combined net profit contribution (net of transaction and loan financing costs) is sufficient in relation to the MDB's return on capital employed target.	NC	NC	NC	NR	NC+	NC	NC	NR	0	0/6
4.5	***Indicators and benchmarks for MDB's additionality*** *(Out of 2x8 = 16)*	0	1	1	2	2	0 (2)	0	0	6 (8)	38% (50%)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
4.5.1	The rating for the MDB's additionality takes into consideration four indicators: • Terms. Would the client have been able to obtain sufficient financing from private sources on appropriate terms? Judgments on this indicator consider pricing (including additional costs arising from MDB conditions that would not be imposed by private investor), tenor, grace period, currency, and timeliness, i.e., the availability of financing without unduly delaying the project. • Was the MDB (because of its being an MDB) needed to reduce the risks or provide comfort (i.e., improve the investors' perceptions of the risks involved) and, thus, to encourage the investors and lenders to proceed? • Was the MDB needed to bring about a fair, efficient allocation of risks and responsibilities, e.g., between the public sector and the private investors? • Did the MDB, improve the venture's design or functioning--in business, developmental, transition, social or environmental terms?	NC	MC-	MC	MC	MC	NC (MC)	NC	NC[14]	4 (5)	4/8 (5/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
4.5.2	The project's additionality is rated using benchmarks substantially consistent with the following: • Excellent: It is clear that (a) the project would not have gone ahead without the MDB *or* (b) absent the MDB, (i) it would have entailed a materially unfair or inefficient allocation of risks and responsibilities or (ii) it would have been materially weaker in business, developmental, transition, social or environmental terms. • Satisfactory: It is likely that (a) the project would not have gone ahead without the MDB *or* (b) absent the MDB, (i) it would have entailed an unfair or inefficient allocation of risks and responsibilities or (ii) it would have been weaker in business, developmental, transition, social or environmental terms. • Partly unsatisfactory: It is likely that (a) the project would have gone ahead without the MDB *and* (b) the MDB made no contribution to (i) a fair or efficient allocation of risks and responsibilities and (ii) the business, developmental, transition, social or environmental performance of the project. Unsatisfactory: It is clear that (a) the project would have gone ahead without the MDB *and* (b) the MDB made no contribution to (i) a fair or efficient allocation of risks and responsibilities and (ii) the business, developmental, transition, social or environmental performance of the project.	NC	NC+	NC	MC	MC	NC (MC)	NC	NC[15]	2 (3)	2/8 (3/8)
4.6	***Indicators for MDB's work quality*** *(Out of 4x8 = 32)*	3	4	2	4	3 (4)	1 (4)	4	4	25 29	78% (91%)
4.6.1	The rating for the MDB's work quality is based partly on *at-entry screening, appraisal and structuring work*; i.e., how effectively the MDB carries out its work prior to approval of the investment.	MC	MC	MC	MC	MC	NC (MC)	MC	MC-	7 (8)	7/8 (8/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.6.2	The rating for the MDB's work quality is based partly on its *monitoring and supervision quality*, i.e., how effectively the MDB carries out its work after approval of the investment.	MC	MC	MC	MC	MC	NC (MC)	MC	MC-	7 (8)	7/8 (8/8)
4.6.3	The rating for the MDB's work quality is based partly on its *role and contribution*, i.e., the quality of the MDB's contributions from inception to evaluation. The rating judgment considers compliance with basic operating principles, the MDB's contribution to client capacity building objectives (as relevant), the operation's consistency with furtherance of the MDB's corporate, country and sector strategies, and its clients' satisfaction with the MDB's service quality.	NC	MC-	NC	MC	NC (MC)	NC (MC)	MC	MC	4 (6)	4/8 (6/8)
4.6.4	Assessments of the MDB's work quality should be made independently of the ratings assigned for development or transition outcomes and MDB's investment profitability. These assessments, which are benchmarked against corporate good practice, reflect the quality of the MDB's contributions to good or bad outcomes, not the good or bad outcomes themselves.	MC	MC	NC+	MC	MC	MC	MC	MC	7	7/8
4.7	***Standard XASR attachments*** *(Out of 2x8 = 16)*	0	0	1 (2)	0 (2)	0 (2)	0	2	2	5 (10)	31% (63%)
4.7.1	The XASR includes an attachment providing details supporting the financial and economic rate of return estimates (with transparent assumptions and cash flow statements). This attachment provides the basis for review and independent verification of the XASR's judgments and conclusions.	NC	NC+	NC (MC)	NC (MC)	NC (MC)	NC	MC	MC	2 (5)	2/8 (5/8)

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)
(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
4.7.2	The XASR includes an attachment providing, for each safeguard dimension addressed in the MDB's environmental and social guidelines, a comprehensive summary of environmental, worker health and safety, and social outcome compliance information with sufficient evidence from a field visit and/or client reporting to support the assigned outcome and related MDB work quality ratings. This attachment provides the basis for review and independent verification of the XASR's judgments and conclusions.	NC	NC	MC	NC (MC)	NC (MC)	NC	MC-	MC-	3 (5)	3/8 (5/8)
5	**Annual Reporting and Process Transparency** (Out of 54)	**0**	**0**	**7**	**5**	**7**	**5**	**7**	**2**	**33**	**61%**
5.1	***Annual synthesis reporting: Annual Review*** *(Out of (5x8) + (1x7) = 47)*	0	0	6	5	6	5	6	1	29	62%
5.1.1	CED prepares an annual review addressed to the MDB's management, staff and Board of Directors. The scope of the annual review includes, *inter alia*, a synthesis of the CED's validated findings from all XASRs and PERs generated and reviewed during the period covered.	NC	NC+	MC	MC	MC	MC	MC	NC+	5	5/8

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
5.1.2	The annual review should provide sufficient information to make the reader aware of possible biases in the sample of projects covered by the annual review. Consequently, the annual review: • Describes how the population was identified and how the sample was selected. If stratification was applied or part of the sampling was non-random, the annual review states the rationale. • Reports on the number of XASRs and PERs for the year. • Includes an annex profiling the important characteristics of the evaluated sample (e.g., sector, investment size, and percentage of operations affected by specific loss provisions) against the population. • Confirms that all projects selected for evaluation had reached early operating maturity, as defined in the GPS, by the time of their evaluation. • If less than 100% of the population has been covered, provides information on statistical confidence levels and states explicitly whether reported success rates can be attributed to the population.	NC	NC	MC	NC+	MC	NC +	MC	NC	3	3/8
5.1.3	The annual review either (i) describes the ratings criteria and benchmarks in an annex or else (ii) refers to a website providing this information.	NC	NC	MC	MC	MC	MC	MC	NC+	5	5/8
5.1.4	The ratings reported should be those of CED.	NC	NC	MC	MC	MC	MC	MC	MC	6	6/8

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	***Standard***	***AfDB***	***AsDB***	***EBRD***	***IDB-SCF***	***IIC***	***MIF***	***IFC***	***MIGA***	***Total***	***Proportion MC***
5.1.5	*Good practice.* CEDs should disclose the differences between CED and operating staff ratings and the materiality of the differences. Where CED ratings are reflected partly in XASR-As and partly in PERs, the CED should disclose the differences between CED and operating staff ratings separately for the XASRs and the PERs. The disclosure is made in global terms, not on a project-by-project basis and is limited to differences in binary outcome, additionality and work quality success ratings.	NC	NC	MC-	MC-	MC	MC	MC-	NR	5	5/7
5.1.6	*Good practice:* For each rating dimension and indicator, the annual review shows the *proportion of the evaluated sample* in each performance-rating category.	NC	NC	MC	MC	MC	MC	MC	NC+	5	5/8
	Best practice: The annual review also shows, by dimension, the *proportion of total disbursed MDB financing* for the sample that is in each performance-rating category.	NC	NC	MC	NC	MC	NC	MC	NC	3	3/8
5.2	***Process transparency: Annual Report*** *(Out of 1x7 = 7)*	0	0	1	0	1	0	1	1	**4**	**57%**
5.2.1	CED reports periodically (at least every three years) to the MDB's management and Board on the quality and efficacy of the MDB's evaluation system, including the self-evaluation system, any gaps in coverage of the MDB's operations, the work of CED, and the generation and application of lessons learned in new operations. In addition, it submits to the MDB's management and Board the periodic benchmarking reviews of the consistency of the MDB's practices with the GPS.	NC	NC+	MC	NC+	MC-	NR	MC	MC-	4	4/7

Annex 1

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)

(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

Std. #	*Standard*	*AfDB*	*AsDB*	*EBRD*	*IDB-SCF*	*IIC*	*MIF*	*IFC*	*MIGA*	*Total*	*Proportion MC*
6	**Identification of Lessons, Dissemination, and Ensuring Application of Lessons** (Out of 0)	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Grand Totals • ***Number of MC Ratings*** ○ ***Adoption*** ○ ***Adoption & Correct Application*** • ***Possible Number of MC Ratings***	**25** **24** **54**	**35** **34** **54**	**40** **39** **48**	**46** **38** **51**	**51** **41** **52**	**36** **23** **48**	**50** **50** **54**	**37** **37**[16] **51**	**320** **286** **412**	
Percentage of MC Ratings • ***Adoption*** • ***Adoption & Correct Application***	**46** **44**	**65** **63**	**83** **81**	**90** **75**	**98** **79**	**75** **48**	**93** **93**	**73** **73**[17]		**78** **69**

FREQUENCY DISTRIBUTION OF MATERIALLY-CONSISTENT RATINGS FOR GOOD PRACTICE HARMONIZATION STANDARDS, BY MEMBER BASED ON ADOPTION & SUBSTANTIALLY FULL IMPLEMENTATION OF STANDARDS AND, WHERE DIFFERENT, JUST ADOPTION OF STANDARDS (BETWEEN PARENTHESES)
(EXCLUDING EXPERIMENTAL STANDARDS, EXPERIMENTAL COMPONENTS OF STANDARDS, & "OTHER" STANDARDS)

ENDNOTES

[1] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA notes that it cannot comply with either the good practice or best practice standards (requiring statistical significance at the 95% confidence level). This would imply a near 100% coverage of the population of MIGA projects underwritten.

Two other members are preparing XASRs on 100% of their operations.

[2] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA notes that MIGA does not make investments, and that insurers measure profitability on a portfolio basis, not on the basis of an individual guarantee contract. This part of the standard does not apply to MIGA, and should be (sub-) rated NR.

MIGA's Guidelines for Preparing a Project Evaluation Report for Non-Financial Sector Projects suggest that MIGA has accepted the principle of considering contributions to MIGA's profitability but that it has not yet made a firm decision on how to do so, The Guidelines call for evaluation reports to rate, *inter alia*, "the guarantee project's contribution to MIGA's profitability." The Guidelines, however, indicate that the chapter on the project's contribution to MIGA's financial results is "To be finalized." In any event, this review did not "(sub-)" rate individual components of standards.

[3] MIGA's evaluation unit has requested that the following note be added:

> The description of "MDB's work quality" in parenthesis should also refer to *"Quality of Underwriting."*

[4] MIGA's evaluation unit has requested that the following note be added:

> For the reasons below, 4.1.1 should be rated **MC-** (a) MIGA notes that since MIGA does not make "investments," "investment profitability" does not apply to MIGA; moreover, insurers measure profitability on a portfolio basis, not on the basis of an individual insurance contract. Thus this requirement of the GPS does not apply to MIGA (NR)**.** (b) MIGA notes that while MIGA does not have an indicator actually called "Additionality", in substance, this aspect is rated and reported as *"MIGA's Role and Contribution"*. IEG-MIGA requests that rating 4.1.1 be corrected before issuing the Report.

With respect to (a), see endnote 2, above. With respect to (b), the WGPSE specifically decided at its March 14, 2005, meeting that old GPS 26 should be "Split to separate additionality from work quality…." Notwithstanding this decision, reflected in the GPS, MIGA has continued to include additionality with role and contribution. See also my comments on footnote 21 to the text of the report.

[5] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to this rating and does not accept it. IEG-MIGA notes that MIGA was rated "MC-" on this GPS in the previous Matrix (February 8, 2010) circulated by the consultant. IEG-MIGA meets this standard in all components, as evidenced from the sample of PERs reviewed by the consultant. IEG-MIGA requests this rating be corrected to **MC-** prior to the issuance of the Report.

The GPS calls for the rating to be based "secondarily" on "(i) the project's contributions to other business goals articulated at approval and (ii) the project *company's* prospects for sustainability and growth." MIGA's Guidelines call for assessing performance relative to objectives *only* where no relevant quantitative information is available and do not call for assessing the company's prospects for sustainability and growth.

[6] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC+ rating**.** For the reasons below, 4.3.3 should be rated **NR:** (a) IEG-MIGA notes that MIGA as an insurance agency does not require the information needed to calculate a WACC during guarantee underwriting. (b) As a guarantor of an investor's loan or equity in a project ("project company" in MIGA's terminology), MIGA would not normally have access to the information needed to calculate the company's weighted average cost of capital (WACC). (c) This standard does not apply to MIGA as an insurer. IEG-MIGA requests this be corrected before issuing the Report. .

The WGPSE agreed on the requirements for a not relevant rating (Table 2 of report). MIGA does not meet these requirements with respect to GPS 4.3.3.

[7] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC+ rating. For the reasons below, 4.3.4 should be rated **MC-:** (a) IEG-MIGA notes that MIGA has identical benchmarks as IFC (which is rated MC-), except for the use of WACC. (b) As noted above, WACC is not used by MIGA as an insurance agency, and this part of the standard does not apply to MIGA as an insurer. (c) This GPS should be rated consistently across institutions, thus MIGA should be rated the same as IFC, ie, **MC-.** IEG-MIGA requests rating 4.3.4 be corrected before issuing the Report.

IFC received an MC- rating because its benchmark for a satisfactory rating (the key rating in IFC's reports and in comparisons among members) was fully consistent with the GPS. For IFC, a satisfactory rating requires that the FRR be equal to or greater than the weighted average cost of capital. MIGA, in contrast, uses an arbitrary hurdle rate of 10%, which could be greater than, equal to, or less than the weighted average cost of capital.

---

[8] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC rating. For the reasons below, 4.3.5 should be rated NR: (a) IEG-MIGA notes that MIGA as an insurance agency does not require the information needed to calculate the WACC specified in GPS 4.3.3, which GPS 4.3.5 refers to, during guarantee underwriting. (b) As a guarantor of an investor's loan or equity in a project ("project company" in MIGA's terminology), MIGA would not normally have access to the information needed to calculate the company's weighted average cost of capital (WACC) specified in GPS 4.3.3, which GPS 4.3.5 refers to. (c) This standard does not apply to MIGA as an insurer, since GPS 4.3.3 does not apply to MIGA, which GPS 4.3.5 refers to. IEG-MIGA requests that rating 4.3.5. be corrected before issuing the Report. IEG-MIGA also notes that MIGA was not included in WGPSE's 2006 consensus-based development of the GPS3 where this rating was agreed; and MIGA disagrees with this standard.

See earlier comments on WACC and, with respect to MIGA's participation in the WGPSE discussions that resulted in the approval of the Third Edition, see footnote 3 in the present report.

[9] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC+ rating. For the reasons below, 4.4.1 should be rated **NR:** (a) MIGA notes that since MIGA does not make "investments," "investment profitability" does not apply to MIGA; (b) moreover, insurers measure profitability on a portfolio basis, not on the basis of an individual insurance contract. Thus this requirement of the GPS does not apply to MIGA and should be rated **NR.**

See comments on earlier MIGA notes.

[10] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC+ rating. For the reasons below, 4.4.1 should be rated NR: (a) MIGA notes that since MIGA does not make "Investments," "investment profitability" does not apply to MIGA; (b) moreover, insurers measure profitability on a portfolio basis, not on the basis of an individual insurance contract. Thus this GPS does not apply to MIGA and should be rated NR.

See comments on earlier MIGA notes.

[11] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC rating. For the reasons below, 4.4.1 should be rated **NR:** (a) MIGA notes that since MIGA does not make "investments," "investment profitability" does not apply to MIGA; (b) moreover, insurers measure profitability on a portfolio basis, not on the basis of an individual insurance contract. Thus this GPS does not apply to MIGA and should be rated **NR.**

See comments on earlier MIGA notes.

[12] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA objects to the NC rating. For the reasons below, 4.4.2 should be rated **NR:** (a) MIGA notes that this standard does not apply to MIGA, as MIGA itself does not articulate, specify or in any other way indicate any expectations with respect to guarantee performance at approval. This is because insurers measure profitability on a portfolio basis, not on the basis of an individual insurance contract. Thus this requirement of the GPS does not apply to MIGA and should be rated **NR.**

See comments on earlier MIGA notes.

[13] MIGA's evaluation unit has requested that the following note be added:

> MIGA does not make loans. The rating on GPS 4.4.2 is not relevant to MIGA and should be rated **NR.**

Rating corrected. (But best-practice ratings not taken into consideration in scoring performance.)

[14] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA disagrees with the overall NC rating on standard 4.5.1 and believes it should be rated **MC** for the following reasons: (a) IEG-MIGA is fully compliant on the 1st and 2nd dimensions of the additionality measure that are applicable to MIGA. (b) The 3rd and 4th dimensions of standard 4.5.1. do not apply to MIGA, and should be **(sub-) rated NR**: Since insurers are approached for coverage only after a transaction has already been designed and just before financial closure, the 3rd dimension – *"Was the MDB needed to bring about a fair, efficient allocation of risks and responsibilities, e.g., between the public sector and the private investors?"* cannot not apply to MIGA and should be rated NR, not NC. For the same reasons, the 4th dimension – *"Did the MDB, improve the venture's design or functioning--in business, developmental, transition,... terms?"* does not apply to MIGA. (c) While MIGA does not have a "fourth" overall indicator called "Additionality", the additionality aspect is assessed, rated and reported as an indicator called "*MIGA's Role and Contribution"*.

See earlier comments on separating additionality from role and contribution and on definition of "not relevant" agreed by the WGPSE.

[15] MIGA's evaluation unit has requested that the following note be added:

> IEG-MIGA disagrees with the overall NC rating on standard 4.5.2 and believes it should be rated **MC** for the following reasons: (a) MIGA's rating benchmarks for "*MIGA's Role and Contribution"* are materially consistent with these benchmarks, except thsat they do not refer to the items listed in (b) which do not apply to MIGA, for the reasons explained in MIGA's Comment (see footnote) on 4.5.1. (b) While MIGA does not have a "fourth" overall indicator called "Additionality", the additionality aspect is assessed, rated and reported as an indicator called "*MIGA's Role and Contribution"*.

See earlier comments on MIGA's notes.

[16] MIGA's evaluation unit has requested that the following note be added:

> On the basis of the explanations given in MIGA's footnoted Comments on GPS standards #4.1.1, #4.3.2, #4.3.4, #4.5.1, and #4.5.2, MIGA believes the number of MC ratings (adoption and correct application) should be 42.

See earlier comments on MIGA's notes.

[17] MIGA's evaluation unit has requested that the following note be added:

> On the basis of the explanations given in MIGA's footnoted Comments on GPS standards #4.1.1, #4.3.2, #4.3.4, #4.5.1, and #4.5.2, MIGA believes the percentage of MC ratings (adoption and correct application) should be 82 %.

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
1		**Roles of Independent and Self-Evaluation**	
1.1		***Structure and role of independent evaluation, i.e., of Central Evaluation Department (CED):***	
1.1.1	Harmonization	CED has a Board-approved mandate statement, designed to ensure independence and relevance.	Mandate must be approved by Board of Directors to be rated as MC.
1.1.2	Harmonization	The mandate provides that the Board of Directors oversees CED's work and that the CED's reporting line, staff, budget and functions are organizationally independent from the MDB's operational, policy and strategy departments and related decision-making. In addition, the mandate specifies that the Board has the ultimate decision authority for (1) hiring and terminating CED head and staff; (2) CED head's appointment terms and reporting structure; (3) CED head's and staff's grading, performance reviews and pay increases; and (4) the CED's budget.	MC rating requires explicit statements in mandate or equivalent document. I have, nevertheless, not required reference to performance reviews, since grading and pay increases are what ultimately counts.
1.1.3	Other	CED operates with full autonomy but in close consultation with the MDB's other departments to ensure as far as possible (subject to the primacy of sound evaluative principles and practices) coherence of corporate standards (as among operations, portfolio and strategy analysis, and evaluation) and good prospects for corporate ownership of CED's findings and recommendations for improvement. To this end, CED seeks alignment, as far as possible, of performance measures and standards used in evaluations and in non-CED reports to management and Board.	Critical word is "operates." Does not require explicit statement in mandate. This GPS deals with actions taken by CED. GPS 5.4.3 deals with outcomes.
1.1.4	Harmonization	Under its mandate, CED's scope of responsibility extends, without restriction, to all determinants of the MDB's operational results.	If management has right to approve or disapprove program, it can limit scope of responsibility.
1.1.5	Harmonization	The mandate states that CED has unrestricted access to MDB's staff, records, co-financiers, clients and projects. The mandate may, however, allow for restrictions on access to clients and projects in jeopardy cases, where an evaluator's visit could prejudice the MDB's financial interests or materially increase the risk of litigation. Should client access be restricted in jeopardy cases, the number of such cases should be reported in the MDB's annual report or annual review.	MC rating requires explicit statement in mandate, covering among other things access to clients and projects. I have not, however, insisted on explicit statement about co-financiers.
1.1.6	Harmonization	The mandate provides that CED transmits its reports to MDB's Board after review and comment by management but without management clearance or any management-imposed	Essence of standard is freedom from management clearance or any management-imposed restrictions on

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		restrictions on their scope and contents.	scope and contents of CED reports. MC rating requires explicit statement in mandate providing for CED to be able transmit its reports without management clearance or any management-imposed restrictions on their scope and contents.
1.1.7	Other	The mandate provides that CED's manager holds grade-rank at least equal to that of operational department directors.	MC requires that provision be embodied in mandate. Specification of title meets this requirement if title implies grade-rank equivalent to operational department directors.
1.1.8	Other	The CED or the MDB has issued guidelines, applicable to all CED managers and staff members, designed to ensure that CED evaluations are—and are perceived to be—devoid of any conflict of interest.	
1.2		***Responsibilities of operations departments in self-evaluation:***	
1.2.1	Harmonization	*Good practice.* Execute XASRs on investments selected pursuant to GPS 2.3.1 and 2.3.2 in accordance with CED’s sample selection and evaluation guidelines. *Best practice.* In addition, deliver XASRs according to a schedule designed to spread the review load throughout the program year and allow CED to complete the annual review on schedule.	Applicable only to institutions that prepare XASRs, i.e., that do not limit evaluations to PERs. **Add after “investments” here *et passim* “(including loans and guarantees).”**
2		**Evaluation Timing, Population, Coverage and Sampling**	
2.1		***Identification of population from which sample for evaluation is to be drawn; timing of consideration for evaluation:***	
2.1.1	Harmonization	Taking into consideration information on project maturity status provided by other departments, CED determines the *population* from which the investments to be evaluated each year are to be drawn.	Similar to GPS 2.3.1, which calls for CED’s selecting the *sample* of projects to be evaluated. The key point in both standards is the role of CED.
2.1.2	Harmonization	The population from which the investments to be evaluated each year are to be drawn consists of the investments that will have reached *early operating maturity* (as defined in GPS 2.1.3 and 2.1.4) during the year.	For guarantee operations, references to “disbursed investments,” here *et passim*, should be read as “committed guarantees.”

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		• Subject to certain exclusions, specified below, the population includes all disbursed (including partially cancelled) investments--whether still active or already closed (paid-off, sold or written off)--that have reached early operating maturity. The population also includes investments already closed, even if they never reached early operating maturity. • Excluded from the population are dropped and cancelled investments, very small investments made under special promotional programs (e.g., direct investments in small businesses that are evaluated on a program basis through a CED special study), subscribed rights offerings and investments undertaken to help finance cost overruns on projects previously financed by the MDB. • Projects that have not yet reached early operating maturity are excluded from the current evaluation year's population and rolled forward for inclusion in the population in a future year when they will have reached early operating maturity. • Investments are included in the population from which the sample for evaluation is drawn only once, i.e., only for the year in which they will have reached early operating maturity	Since visits to closed investments may not be feasible and since operational staff may be unwilling to devote resources to visiting closed investments, CED may carry out abbreviated desk reviews to evaluate these operations. The important thing is that they not be excluded from the population, which would introduce bias in reporting on overall outcomes.) For already-closed investments that are selected for an XASR, the XASR consists of the last available supervision report and the attached evaluative addendum. Based on what is needed to achieve the objective of comparability and the wording of GPS 5.1.2, I have ignored "during the year" in assigning ratings for this standard. I have, nevertheless, still assigned NC ratings when members have evaluated projects that have not yet achieved early operating maturity or that otherwise fall short of the standard. **The 4th edition of the GPS should clarify this standard and resolve the apparent inconsistency with GPS 5.1.2. The issue is whether projects should be included in the population *only* in the year they reach early operating maturity or whether they can *instead* be included in the population of a subsequent year. (Some projects have been evaluated four or more years after reaching early operating maturity.) To ensure that evaluation of unsuccessful investments is not unduly deferred and to ensure that successful projects are not included in the population more than once (which would increase the probability of their being selected for evaluation), the GPS should specify that all projects are to be included in the population in the year they reach early operating maturity or, when the information needed to determine whether the project would reach early operating maturity during the year was not available on time, during**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
			the following year. **The second sentence of the first bullet needs clarification. Two members misunderstood it. The standard should make clear that (i) investments already closed should be treated the same way as other investments and (ii) investments unlikely to reach early operating maturity should be evaluated within a year of the time when it is determined that they are unlikely to reach early operating maturity. The second point should apply to all projects.**
2.1.3	**Harmonization**	All operations other than the financial markets operations specifically covered by GPS 2.1.4 are deemed to have reached *early operating maturity* when (a) the project financed will have been substantially completed, (b) the project financed will have generated at least 18 months of operating revenues for the company and (c) the MDB will have received at least one set of audited annual financial statements covering at least 12 months of operating revenues generated by the project.	MC requires that definition of early operating maturity be consistent with GPS. *All* tests must be met for MC rating. **An additional standard may be needed for operations not covered by GPS 2.1.3 or 2.1.4, i.e., projects where the concept of project completion is not relevant and the principal objective is not to assist identifiable capital expenditure projects or sub-projects, e.g., trade finance facilities, facilities for warehousing of financial instruments prior to securitization, corporate financing, etc.**
2.1.4	**Harmonization**	Financial markets projects where the principle objective is to assist identifiable capital expenditure sub-projects (rather than to contribute to institutional development or institution building) are deemed to have reached *early operating maturity* after the elapse of at least 30 months following the MDB's final material disbursement for sub-loans or sub-investments, i.e., ignoring disbursements for small follow-up investments in existing client companies and disbursements to cover management fees or other expenses of investment funds.	Applicable only to institutions that provide financing to intermediaries or investment funds where the principle objective is to assist identifiable capital expenditure sub-projects. MC requires that definition of early operating maturity be consistent with GPS. **Wording should be changed to say "identifiable capital expenditure sub-projects or identifiable equity investments."** **Wording needs to be changed for MIGA, which does not make disbursements for sub-loans.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
2.2		***Evaluation coverage:***	
2.2.1	**Harmonization**	*Good practice:* Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample of sufficient size to establish, *for a combined three-year rolling sample,* success rates at the 95% confidence level, with sampling error not exceeding ±5 percentage points, for the population's development (transition) outcome, MDB investment outcome, additionality and MDB work quality. *Transitional good practice:* Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample equivalent to 60% or more of the investments in the population. In using this standard, an MDB reports on the confidence level and sampling error applicable to the success rates for the population's development or transition outcome, MDB investment outcome, additionality and MDB work quality. An MDB can use this standard only until its combined three-year rolling population of projects reaching early operating maturity reaches 50. *Best practice-Alternative 1*: Preparation of XASRs (with XASR-As), PERs, or a combination of the two on a random, representative sample of sufficient size to establish, *for a combined three-year rolling sample,* success rates at the 95% confidence level, with sampling error not exceeding ±5 percentage points, for the population's development or transition outcome, MDB investment outcome, additionality and MDB work quality *within each of the MDB's current strategically targeted groups.* *Best practice-Alternative 2*: Preparation of XASRs (with XASR-As), PERs, or a combination of the two on 100% of the investments in the population.	The standard calls for meeting certain confidence level and sampling error tests. These tests would need to be applied to the combined sample for the three years ending with the most recent year for which evaluation results are available. Thus, in an MDB's annual review for, say, 2006, it would look at the combined results of the evaluations carried out during 2003, 2004 and 2005. An MC rating requires that a Member's practices be consistent with the good practice standard, the transitional good practice standard, or one of the best practice standards. **The redundancy of specifying a confidence interval of 95% and a sampling error not exceeding ±5% should be eliminated.**
2.3		***Sampling:***	
2.3.1	Harmonization	The CED selects the operations for XASRs and PERs from the evaluation year's population (as defined above), subject to the following standard.	Similar to GPS 2.1.1, which calls for CED to determine the *population* from which the sample is to be drawn. The key point in both standards is that the CED selects the sample, not management or the operational staff.
2.3.2	Harmonization	If coverage is less than 100%, the sample should be both random and representative. Notwithstanding this principle, a CED may wish to select projects to be covered by PERs	Only random or stratified random samples support performance inferences about the sampled population. Representativeness is important for *prima facie*

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		based, e.g., on the potential for learning, the high profile of an operation, credit and other risks, whether the sector is a new one for the MDB, the likelihood of replication, or the desirability of balanced country and sector coverage. If so and if the CED wishes to combine the PER with the XASR-A findings in reporting annual success rates, it uses stratified sampling methodology, as follows: • It splits the population into two strata. The first consists of the projects CED selects for PERs. The second consists of the remaining projects, i.e., the population other than the projects selected for the first stratum. • CED evaluates 100% of the first stratum. The success rates from the sample are, thus, identical to the success rates for this portion of the population. • CED selects a random sample from the second stratum. The operational staff prepares XASRs on the projects selected, and CED prepares XASR-As on these projects (or a random sample of them). The sample is sufficiently large to give reliable estimates of the success rates for that stratum. • Based on the weight of each stratum in the overall population, CED then calculates the weighted average success rates and sampling errors, following the normal procedures for stratified sampling. If the CED wishes to select projects to be covered by PERs as above *but does not wish to combine the PER with the XASR-A findings in reporting annual success rates*, it draws the sample to be covered by XASRs from the full population for the year, without previously eliminating the projects to be covered by PERs. To the extent that specific projects may be selected for XASR-As and PERs, CED would use the PER ratings, rather than the XASR-A ratings in reporting on success rates, since CED will have carried out a more rigorous review in these cases.	plausibility of the results.
3		**Guidelines, Execution and Validation**	
3.1		***Guidelines & familiarisation:***	
3.1.1	Harmonization	In consultation with operations departments, CED prepares, refines and disseminates guidelines for the preparation of XASRs and PERs in sufficient detail to promote consistency and objectivity in execution scope, analysis and ratings. As part of dissemination efforts, some CEDs may wish to conduct workshops to familiarize the XASR teams with requirements and supporting documentation for achieving good-	

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		practice execution.	
3.1.2	Harmonization	*Good practice*: The guidelines include ratings guidelines with benchmarks and standard reporting templates that include the performance ratings matrix. *Best practice*: The guidelines also include related documentation, such as an overview of the XASR program, a description of efficacious execution process steps, good-practice examples of XASRs from previous years' samples, and a list of execution mistakes to avoid (informed by past XASRs).	
3.1.3	Other	CED maintains these guidelines on its website.	
3.2		***Execution:***	
3.2.1	Harmonization	The research for XASRs and PERs draws from a file review; discussions with available staff involved with the operation since its inception; independent research (e.g. on market prospects); a field visit to obtain company managers' insights and to the project site to observe and assess outcomes; and discussions with parties who are knowledgeable about the country, company and project (e.g. MDB specialists, company employees and auditors, suppliers, customers, competitors, bankers, any relevant government officials, industry associations, and local NGOs).	MC requires, *inter alia*, a field visit for substantially all XASRs and PERs.
3.3		***Review and independent validation:***	
3.3.1	Other	*Good practice.* The standard transmittal memo on the XASRs executed by operations department staff incorporates the approval (or electronic check-off) by the responsible operations department manager. *Best-practice.* In addition, the standard transmittal memo on the XASRs executed by operations department staff incorporates the approval (or electronic check-off) and, if relevant, cites disagreements by other departments, e.g., technical, environmental, economics and syndications.	Applicable only to institutions that prepare XASRs, i.e., that do not limit evaluations to PERs. The XASR findings comprise a set of representations by management (through the CED) to the Board, and a sign-off or check-off comprises the only written evidence of the operating management's endorsement of the staff's representations.
3.3.2	Other	To provide transparency with respect to field visits (GPS 3.2.1), the XASR or PER or the XASR transmittal memo provides information on when field visit took place and who (i.e., representatives of which departments) participated in the field visit.	This information can be included in the transmittal memo or in the XASR or PER. Some MDBs may wish to call for information on persons interviewed (with titles and affiliations).
3.3.3	Harmonization	CED conducts an independent review of each XASR to verify scope responsiveness, evident	Applicable only to institutions that prepare XASRs,

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		reliability of the analysis, impartiality and consistency in ratings judgments, and appropriateness and completeness of the identified lessons, and then, *for each randomly selected XASR to be used in the annual synthesis report on evaluation results*, prepares an XASR-A on the final-edition XASR that records its independent judgments on the report's quality in relation to the guidelines, assigned ratings and lessons.	i.e., that do not limit evaluations to PERs.
3.3.4	Other	*Best-practice*: For XASRs recommended by CED, the relevant vice president, central portfolio manager, credit manager, or other manager at a level higher than the responsible officer and his or her manager chairs a review meeting that is attended by the XASR team and their managers, CED, and representatives of specialist departments (e.g. credit, technical and environmental, economics, legal, syndications and special operations) as relevant. Operations staff responsible for the operation at entry are invited to attend the review meeting, comment on the XASR's findings, or both.	
3.3.5	Harmonization	Following preparation of each draft XASR-A, CED reviews with the XASR team and its manager the basis for its judgments where its ratings differ from those in the final edition XASR.	Applicable only to institutions that prepare XASRs, i.e., that do not limit evaluations to PERs.
3.3.6	Other	At the end of the program year and prior to submitting its annual review, CED sends a ratings validation variance memo to the responsible senior operations manager, with copies to the relevant XASR teams and their managers.	Applicable only to institutions that prepare XASRs, i.e., that do not limit evaluations to PERs.
4		**Evaluative Scope**	
4.1		***Performance dimensions evaluated:***	
4.1.1	Harmonization	*Good practice*: The scope of the XASR (and XASR-A) or PER includes, at a minimum, • The *development or transition outcome of the project and, to the extent provided in GPS 4.3.3-4.3.7 and GPS 4.3.17, the project company*, i.e., the "results on the ground" relative to the MDB's mission. • The *MDB investment's profitability* (contribution to its corporate profitability objective), • The *MDB's additionality, and* • The *MDB's work quality* (also referred to as bank handling, operational effectiveness, or execution quality).	An MC rating requires that additionality be rated separately. Focusing on the underlying objective of fostering comparability among Members, I have not insisted that additionality be covered by XASRs and PERs, but I have required that XASR-As treat additionality as a separate dimension, that additionality not be taken into consideration in the XASR-As' work quality ratings, and that the CED's annual reviews report on the additionality ratings and the work quality ratings shown in the XASR-As (i.e., based on the GPS).

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
			First bullet should refer to GPS 4.3.16, rather than 4.3.17. **To eliminate future disputes, the text should make clear that ratings need to be assigned for each dimension.**
4.1.2	Harmonization	The operation's performance under each of these dimensions is analyzed according to standard indicators, and the operation's performance for each indicator is rated according to criteria and benchmarks specified in the guidelines.	
4.1.3	Harmonization	The performance reflected in the relevant indicator ratings is synthesized into ratings for each of the four performance dimensions, specified above.	Some WGPSE members consider that an overall rating, synthesizing the ratings for the four dimensions, would be useful. They indicate that their Boards of Directors want an overall judgment on the operations evaluated. Others disagree. They argue that seeking to provide an overall rating for the results on the ground for the host country (an outcome), the contribution to the MDB's profitability (another outcome), work quality (an input that may or may not contribute to these outcomes), and additionality would yield a rating that has no clear meaning. They argue further that their Boards of Directors primarily want to know whether the operations financed are contributing to the institution's objectives and that Boards members recognize that attribution of these results to the institution's contributions or to other factors is a separate question. Although the present standards do not bar a member from assigning overall ratings, they do not call for such ratings. Members wishing to experiment with such ratings may do so. **This issue can be reconsidered in developing the 4th edition.** **This standard can probably be combined with GPS 4.1.1.**
4.2		***Performance ratings—principles and benchmarks:***	

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
4.2.1	Harmonization	Within the rating scales (e.g., ranging from *unsatisfactory* to *excellent*), there should be balance between positive and negative characterizations (i.e., if there are four ratings, two are less than good and two are good or better). The words used to describe these ratings should accurately reflect whether the judgments are less than good or else good or better.	
4.2.2	Harmonization	Each of the evaluated performance attributes is assigned a rating using a 4- to 6-point scale for each indicator ratings and a 4- to 6-point scale for each synthesis rating.	
4.2.3	Harmonization	Ratings of non-quantitative indicators require that relative qualitative judgments be made. The criteria should reflect the extent to which performance has been consistent with the MDB's policies, prescribed standards for corporate sustainability and recognized good-practice standards. The criteria for the judgments should be clearly specified in the guidelines for the preparation of XASRs and in the CED's annual review.	**Should make clear that this standard applies also to PERs.**
4.2.4	Harmonization	The synthesis ratings for the four dimensions (development or transition outcomes, profitability to the MDB, the MDB's additionality and the MDB's work quality) reflect summary qualitative performance judgments based on the underlying indicator ratings. They are not simple averages of the indicator ratings.	
4.2.5	Harmonization	Outcomes for each of the indicators are assessed on a "with v. without project" basis.	MC requires that this rule be specifically stated.
4.3		***Indicators and benchmarks for the development or transition outcome (the project's contribution to (a) the company's business success;(b) the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy; and (c) economic development; and (d) the project's overall environmental performance):***	
4.3.1	Harmonization	Assessments of development or transition outcomes for each of the development or transition outcome indicators take into consideration the sustainability of the results.	**This standard might be more appropriately be built into the standards for each of the outcome indicators.**
4.3.2	Harmonization	The project's development or transition outcome is based partly on the *project's contribution to the company's business success*, measured primarily by the real after tax returns and secondarily by (i) the project's contribution to other business goals articulated at approval and (ii) the project *company's* prospects for sustainability and growth.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. **This GPS calls for considering *secondarily* (i) the project's contribution to other business goals articulated at approval and (ii) the project company's prospects for sustainability and growth.**

Std. #	**Type of Std.**	**Standard**	**Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards**
			GPS 4.3.4-4.3.7 calls for taking these factors into consideration "*where appropriate*," and GPS 4.3.3 does not take these factors into consideration. To eliminate repetition and eliminate the possibility of inferring that these factors do not apply to GPS 4.3.3, I suggest dropping GPS 4.3.2 and making clear in GPS 4.3.3-4.3.7 that these factors should be taken into consideration *secondarily*. These factors need to be covered in GPS 4.3.3-4.3.7 because of the need for different wording in some of these standards.
4.3.3	Harmonization	*In rating the business success of capital expenditure projects where the incremental costs and benefits can be quantified,* the project's after-tax financial rate of return in real terms (FRR) is compared with the company's weighted average cost of capital (WACC). Financial and economic theory holds that a firm must expect an after-tax financial rate of return on the funds it invests that is at least sufficient to induce investors to purchase and hold the firm's debt and equity. The investors' return requirements are reflected in the company's WACC, i.e., the weighted average after-tax cost to the company of the yields it must provide on its borrowings and the equity investors' minimally acceptable returns, all adjusted for inflation. More specifically, the WACC is the sum of (a)(i) the average after-tax cost of the company's debt multiplied by (ii) the company's debt as a percentage of its debt plus equity; plus (b)(i) the cost of the company's equity multiplied by (ii) the company's equity as a percentage of its debt plus equity; minus (c) the inflation rate. In estimating the WAAC, the cost of the company's equity is assumed to be the average nominal pre-tax cost of the company's debt plus a 350 bp equity premium.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental with respect to magnitude of equity risk premium, which is to be reconsidered in 2009.* **The standard should make clear that the WACC is to be calculated based on the information available when the investment is made, not on the actual year-by-year figures.** **Before deciding whether to retain the 350 bp equity premium, WGPSE should look at the members' experience with this standard (essentially at IFC, which has the most experience with this standard) to assess whether application of this standard has led to anomalous results. This assessment would address one member's concerns that a standard equity risk premium may be inappropriate in light of differences in ownership structures, sectors, counties, time of investment, etc.** **Also, see GPS 4.3.2.**
4.3.4	Harmonization	The ratings benchmarks for the real after-tax FRR for projects covered by this standard are: • Excellent: FRR exceeds the average nominal cost of the company's borrowings by 700	Applicable only to institutions that seek to contribute to economic development and poverty reduction.

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		bp or more. • Satisfactory: FRR is equal to or greater than the WACC but less than the FRR required for an excellent rating. • Partly unsatisfactory. FRR is equal to or greater than what the WACC would be if the shareholders earned the same return as the lenders (i.e., if the equity premium were zero, rather than 350 bp over the nominal pre-tax cost of the company's debt) but is less than the FRR required for a satisfactory rating. • Unsatisfactory. FRR is lower than the FRR required for a partly satisfactory rating (i.e., the shareholders earned less than the lenders). The business success ratings also take into consideration, where appropriate, (i) the project's contribution to other business goals articulated at approval and (ii) the project *company's* overall prospects for sustainability and growth.	*Experimental with respect to magnitude of equity risk premium and premium for excellent rating, which are to be reconsidered in 2009.* **This standard should be converted into a normal harmonization standard but should be split into a good practice standard, which would specify the benchmark for a satisfactory rating, and a best practice standard, which would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.** **Also, see GPS 4.3.2.**
4.3.5	Harmonization	*In rating the business success of operations involving loans to intermediaries to finance identifiable capital expenditure sub-projects,* the project portfolio's contribution to the after-tax real return on the intermediary's equity is compared with the equity returns implied by the FRR benchmarks calculated using the methodology outlined in GPS 4.3.3). If cost accounting data are available, the information needed to estimate the incremental return on the intermediary's equity can be derived from the intermediary's cost accounting system. When cost accounting data are not available, which will normally be the case, the MDB's or the CED's staff can help the financial intermediary develop rough *ad hoc* estimates of the rate of return. The after-tax real return on equity is calculated from (i) the actual (or typical) spreads and other charges on sub-loans financed by the MDB's loan, (ii) the principal amounts to which these charges would apply each year, (iii) write-offs and expected write-offs (i.e., end of projection period loss provisions) on sub-loans financed by the MDB's loan, (iv) administrative expenses, based on cost accounting data, *ad hoc* estimates, or educated guesses, e.g., as to how the average administrative costs for the sub-loans financed might compare with the institution's overall average administrative expenses, (v) collateral benefits, based on the intermediary's accounting system, *ad hoc* estimates, or educated guesses, (vi) adjustments for inflation and for gains or losses, if any, resulting from exchange rate changes, (vii) income taxes, and (viii) the share of the intermediary's equity needed to back the sub-loans, based either on the intermediary's or the BIS' risk weights. (Exhibit A provides a simplified example of how the return can be calculated.) The assumptions underlying the estimates should be specified in the XASR or an attachment. For projects covered by this standard, the ratings benchmarks for the real equity returns are derived from the benchmarks in GPS 4.3.4 as follows:	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental with respect to magnitude of equity risk premium and premium for excellent rating, which are to be reconsidered in 2009.* **Since two organizations just recently adopted the standard, WGPSE should wait to see whether the standard can be applied and whether it yields reasonable results before deciding what to do.** **In any event, to be consistent with other standards, should split this standard into two standards. One would specify the general principal and methodology. The other would specify the benchmarks.** **The experimental component of this standard should be converted into a normal harmonization standard but should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		• Excellent: (a) The real equity return implied by the FRR required for an excellent rating under GPS 4.3.4. (This benchmark is equivalent to the sum of (i) the WACC (as calculated in accordance with GPS 4.3.4) plus a 250 bp premium (required for an excellent FRR) minus (ii) the average after-tax cost of the company's total debt multiplied by the company's debt as a percentage of its total assets plus (iii) the inflation rate, all divided by (b) the company's equity as a percentage of its total assets.) • Satisfactory: (a) At least (i) the average nominal pre-tax cost of the company's debt plus (ii) 350 basis points but (b) less than the benchmark for an excellent rating. • Partly Unsatisfactory: (a) At least the average nominal cost of the company's debt but (b) less than the benchmark for a satisfactory rating. • Unsatisfactory: Below the average nominal cost of the company's debt. The business success ratings also take into consideration, where appropriate, (i) the project's contribution to other business goals articulated at approval (e.g., reducing the financial intermediary's riskiness) and (ii) the *intermediary's* overall prospects for sustainability and growth.	**See GPS 4.3.2.** **Delete ")" at the end of the first para.**
4.3.6	Harmonization	**[***In rating the business success of investments in funds that finance identifiable equity investments*, the project portfolio's contribution to the fund's real after-tax financial rate of return (FRR) to the investors is compared with the following benchmarks: • Excellent: The project portfolio yields a real after-tax financial rate of return (FRR) to the fund investors that exceeds the compound real annual rate of growth in the Standard & Poor's 500 Index for the corresponding period by 2500 basis points or more. • Satisfactory: The project portfolio yields a real after-tax financial rate of return (FRR) to the fund investors that exceeds the compound real annual rate of growth in the Standard & Poor's 500 Index for the corresponding period by 1500-2499 basis points. • Partly Unsatisfactory: The project portfolio yields a real after-tax financial rate of return (FRR) to the fund investors that exceeds the compound real annual rate of growth in the Standard & Poor's 500 Index for the corresponding period by 500 to 1499 basis points. • Unsatisfactory: The project portfolio yields a real after-tax financial rate of return (FRR) to the fund investors that exceeds the compound real annual rate of growth in the Standard & Poor's 500 Index for the corresponding period by less than 500 basis points. Where appropriate, the business success ratings also take into consideration (i) the *local fund management company's* overall prospects for sustainability and growth and (ii) the project's contribution to other business goals articulated at approval	*Experimental standard. To be reconsidered in 2009.* Applicable only to institutions that seek to contribute to economic development and poverty reduction. **This standard should be reconsidered. The idea of comparing the FRR to an index should probably be dropped and replaced by a standard based on a premium over loan rates when the investment was made. This change would address the criticisms made of this standard and bring the standard into line with the approach used elsewhere in the GPS.** **This standard should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.** **See GPS 4.3.2.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
4.3.7	Harmonization	*In rating the business success of all other projects* (i.e., investments not targeted at specific capital expenditure projects, investments in existing companies where the incremental costs and benefits attributable to the operation cannot be quantified, and investments in financial markets operations that do not finance identifiable capital expenditure sub-projects), the time-adjusted return on invested capital in real terms (ROIC, i.e., the costs and benefits to the company as a whole on a before-after, rather than a with-without, basis) is compared with the company's weighted average cost of capital (WACC). ***The ROIC is calculated only where the CED is convinced that the FRR on the project cannot be calculated.*** The ROIC is based on (i) the initial book value of the company as a whole (including debt and equity), (ii) the annual costs and benefits for the company as a whole, (iii) adjustments for increases or reductions in debt and paid-in share capital, (iv) inflation, and (v) the terminal book value of the company as a whole (including debt and equity). Theoretically, the initial and terminal values should be based on the market value of the company as a whole (including the company's debt and equity), but this information is unlikely to be available for many MDB clients, so the book value is used as a proxy. The ratings benchmarks for the real after-tax ROIC for projects covered by this standard are the same as the benchmarks for capital expenditure projects where the incremental costs and benefits can be quantified, as outlined in GPS 4.3.4. The business success ratings also take into consideration, where appropriate, (i) the project's contribution to other business goals articulated at approval and (ii) the *project company's* overall profitability, adaptability and prospects for sustainability and growth.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental standard. To be reconsidered in 2009.* **This standard should be converted to a normal harmonization standard.** **The GPS should provide detailed instructions for calculating the ROIC. Several members made mistakes in their this regard.** **The standard should make clear that the ROIC is to be calculated after taxes, just as in the case of the FRR.** **See GPS 4.3.2.**
4.3.8	Harmonization	The project's development or transition outcome is based partly on the *project's contributions (positive or negative) to the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy.* In assigning ratings for this standard, the following factors may be considered: • Competition: The project contributes to greater efficiency, quality, innovation or customer orientation of other suppliers through competitive pressures or contributes to restrictions on competition, e.g., by increased protective tariffs, cartels, etc. • Market expansion: Expansion of markets through the project entity's interactions with suppliers (backward linkages) and customers (forward linkages) and through contributions to the integration of economic activities with the national or international economy. • Private ownership and entrepreneurship: Significant increase or consolidation of private provision of goods and services and support for entrepreneurial initiative; or weakening	Ratings for this standard were based on judgments of the extent to which the member's benchmarks were consistent with the GPS. I gave substantial importance to including both positive and negative contributions. **WGPSE should examine in detail the members' experience with projects where normal IRRs could not be estimated (including an assessment of whether application of this standard yielded reasonable results) before deciding on what to do.** **This standard should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	**Type of Std.**	**Standard**	**Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards**
		of support for private ownership and entrepreneurship due to allocation by a financial institution of resources to purchases of government securities or loans to state-owned enterprises. • Frameworks for markets (institutions, laws and policies that promote market functioning and efficiency): Creation or strengthening (or weakening) of public and private institutions that support the efficiency of markets; improvements to (or weakening of) the functioning of regulatory entities and practices; contributions (positive or negative) to government policy formation and commitment, promoting competition, predictability and transparency; contributions to laws that strengthen (or weaken) the private sector and an open economy. • Development of financial institutions and financial markets: Contributions (positive and negative) to the development of sustainable financial institutions and the financial markets in which they operate (including creation of new fund management companies to manage the MDB-supported investment fund, creation of new fund management companies by staff from the management company responsible for the MDB-sponsored fund, and creation of subsequent investment funds); improved financial strength in sector (e.g., by improving asset-liability management); pioneering listing on stock exchange or significant broadening of listed value; first-of-a-kind financial instrument in local market; greater resource mobilization; and improved allocation efficiency. • Transfer and dispersion of skills: Project contributes to significant upgrading of technical and managerial skills beyond the project entity; introduction of new technology or know-how, including financial know-how. • Demonstration effects (spread of new behaviors and activities): Demonstration of replicable products and processes that are new to the economy; new investments stimulated by the project; demonstration of ways of successfully restructuring companies and institutions; demonstration of new ways and instruments to finance private sector activity. • Standards for corporate governance and business conduct: Improvements in (or weakening of) standards, e.g., with respect to accounting standards, disclosure standards, risk management standards and the company's governance quality, reputation and business practices (including corruption) as a positive (or negative) corporate role model and quality investment asset. • Development of physical infrastructure used by other private parties.	**would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.**
4.3.9	Harmonization	The project's contribution to the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy, taking into consideration the project's size, is rated using benchmarks substantially consistent with the following:	*Experimental standard, to be reconsidered in 2009* **The project size is relevant only in defining the benchmark for an excellent rating. For example, if a**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		• Excellent: The project (a) made substantial contributions to the country's private sector development, development of efficient capital markets, or transition to a market economy and (b) had virtually no negative impacts in this respect. • Satisfactory: The project (a) contributed to the country's private sector development, development of efficient capital markets, or transition to a market economy, (b) had a clear preponderance of positive impacts in this respect, but (c) did not meet the requirements for an excellent rating. • Partly unsatisfactory: The project had mainly negative impacts on the country's private sector development, development of efficient capital markets, or transition to a market economy, but the negative impacts are not expected to be of long duration or broad applicability • Unsatisfactory: The project had substantial negative impacts on the country's private sector development, development of efficient capital markets, or transition to a market economy and these impacts are likely to be widespread, of long duration, or both.	**project has a balance of negative impacts, the rating will be negative regardless of the project's size. The point of distinguishing between project size is in judging the positive contribution of, e.g., a major infrastructure project (expected to have a big impact) relative to that of a first leasing company or microfinance institution in a country (which could be a small project but could potentially have huge demonstration effects). The standard should be revised accordingly.** **Revise standard to (a) include idea of sustainability & (b) limit project size to benchmark for excellent rating.** **This standard should be converted into a normal harmonization standard but should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.**
4.3.10	Harmonization	The project's development outcome is based partly on its *contributions to economic development.* Performance is assessed not only by the direct economic costs and benefits to the owners and financiers but also by the economic costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	Applicable only to institutions that seek to contribute to economic development and poverty reduction.
4.3.11	Harmonization	*For capital expenditure projects where the incremental costs and benefits can be separately quantified*, the economic development assessment is based mainly on the project's net quantifiable economic benefits and costs, as measured by the project's real economic rate of return (ERR), but taking into consideration also other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	Applicable only to institutions that seek to contribute to economic development and poverty reduction.
4.3.12	Harmonization	*For loans to intermediaries to finance identifiable capital expenditure sub-project and investments in funds to finance identifiable equity investments,* the economic development assessment is based on (i) the economic contributions of the sub-projects; (ii) the contributions, if any, made to more efficient capital markets; and (iii) other costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental standard, to be reconsidered in 2009.*

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		• Since calculating the overall ERR for the package of sub-projects financed would not be practical, the assessment of the economic contributions of the sub-projects is based on a broad judgment on the range within which the combined ERR would be likely to fall (See GPS 4.3.15). The XASR provides information to substantiate this judgment, including information on portfolio credit or equity performance (e.g., information on non-performing loans, collection rates, write-offs, specific loss provisions, or equity FRRs) and information on distortions that may contribute to a material wedge between financial performance and economic returns.) • The assessment of contributions to more efficient capital markets considers (i) positive contributions, such as reductions in market interest rates attributable to the project or developing the supply of, say, venture capital financing or funding for micro, small or medium enterprises, and (ii) negative contributions, such as encouraging an inefficient allocation of capital because of providing sub-loans at subsidized interest rates. • Other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	**Although this standard needs reconsideration, some rigor needs to be retained. One idea that might explored would be to call for estimating ex-post ERRs on a stratified random sample of sub-projects, designed to cover a disproportionate share of the larger sub-projects.** **Replace "project" by "projects" on second line.**
4.3.13	Harmonization	*For all other projects (* i.e., investments not targeted at specific capital expenditure projects, investments in existing companies where the incremental costs and benefits cannot be separately quantified, and investments in financial markets operations that do not finance identifiable capital expenditure sub-projects), the economic development assessment is based mainly on the net quantifiable economic benefits and costs, as measured by the economic return on invested capital (EROIC), i.e., by the time-adjusted internal rate of return on the economic costs and benefits on a before-after, rather than a with-without, basis but taking into consideration also other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc. ***The EROIC is calculated only where the CED is convinced that the ERR on the project cannot be calculated.*** It is calculated by adjusting the ROIC (GPS 4.3.7) for the factors normally taken into consideration in adjusting the FRR to the ERR, e.g., taxes, subsidies, externalities, etc.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental standard, to be reconsidered in 2009.* **WGPSE should examine in detail the members' experience with projects where normal ERRs could not be estimated (including an assessment of whether application of this standard yielded reasonable results) before deciding on what to do. This standard should be handled the same way as GPS 4.3.8.** **The GPS should provide detailed instructions for calculating the EROIC. Several members made mistakes in their approach.**
4.3.14	Harmonization	*The ratings benchmarks for the ERR or the EROIC for all projects other than financial markets projects with identifiable capital expenditure sub-projects* are as follows: • Excellent if the ERR or the EROIC is 20% or higher.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental standard, to be reconsidered in 2009.*

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		• Satisfactory if the ERR or the EROIC is 10% or higher but less than 20%. • Partly unsatisfactory if the ERR or the EROIC is less than 10% but equal to or greater than 5%. • Unsatisfactory if the ERR or the EROIC is less than 5%. When the other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc. are sufficiently material, the rating may be adjusted upward or downward, particularly when the ERR or EROIC is close to the cut-off points between ratings.	**This standard should be converted into a normal harmonization standard but should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.**
4.3.15	Harmonization	*The ratings benchmarks for the ERR* for financial markets projects with sub-projects are as follows: • Excellent. The evaluation report provides acceptable evidence that (a) the combined ERRs of the sub-projects financed would probably be greater than 20% or (b)(i) the combined ERRs of the sub-projects financed would probably be greater than 10% and (ii) the operation has contributed to the development of a more efficient capital market. • Satisfactory. The evaluation report provides acceptable evidence that (a) the combined ERRs of the sub-projects financed would probably be greater than 10% (but less than 20%) and (b) the operation has not contributed to a less efficient capital market • Partly unsatisfactory. The evaluation report provides acceptable evidence that (a)(i) the combined ERRs of the sub-projects financed would probably be less than 10% (but more than 5%) and (ii) the operation has not contributed to a less efficient capital market, or (b)(i) the combined ERRs of the sub-projects financed would probably be less than 5% but (ii) the operation has contributed to the development of a more efficient capital market. • Unsatisfactory. (a) The combined ERRs of the sub-projects financed would probably be less than 5% or (b) the operation has contributed to a less efficient capital market. When the other material, documented costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc. are sufficiently material, the rating may be adjusted upward or downward, particularly when the ERR or EROIC is close to the cut-off points between ratings.	Applicable only to institutions that seek to contribute to economic development and poverty reduction. *Experimental standard, to be reconsidered in 2009.* **Reconsider in context of revisions of GPS 4.3.12.** **This standard should be converted into a normal harmonization standard but should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.** **Replace "or" by "and" on penultimate line of fourth bullet para.**
4.3.16	Harmonization	The project's development or transition outcome is based partly on the company's overall environmental performance at the time of evaluation. The assessment is based primarily on the MDB's specified standards in effect at investment approval (though compliance with the standards specified at the time of the evaluation may be taken into consideration in assigning	Based on the WGPSE's intentions in approving the 3d edition, I have assigned an MC rating when a member's guidelines call for looking beyond the environmental performance of just the project financed, e.g., by

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		"outstanding" or "excellent" ratings). "Environment" includes the physical environment and, to the extent covered by the MDB's policies, also includes energy efficiency; the quality of the client's environmental management activities; social, cultural, health and safety impacts; and the extent of public consultation and participation.	calling for assessing environmental impacts in the vicinity or area of influence of the project, but not to require the assessment of the environmental performance of the company as a whole. **The wording of this standard should be clarified by adding "in the area of influence of the project" or alternatively "in the vicinity of the project" on line 2 after "performance."** **Delete "partly" on 1st line.**
4.3.17	Harmonization	The environmental performance is rated based substantially on the following benchmarks: For non-financial markets projects: • Excellent: The company (a) meets (i) MDB's at approval requirements (including implementation of the environmental action program, if any) and (ii) MDB's at evaluation requirements; and (b) has either (i) gone beyond the expectations of the environmental action plan or (ii) materially improved its overall environmental performance (e.g., through addressing pre-existing environmental issues) or (iii) contributed to a material improvement in the environmental performance of local companies (e.g., by raising industry standards, acting as a good practice example, etc.) • Satisfactory: The company is in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any). • Partly unsatisfactory: (a) The company is not in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any) but is addressing deficiencies through on-going or planned actions; or (b) earlier non-compliance (subsequently corrected) resulted in environmental damage that has not been corrected. • Unsatisfactory: (a) The company is not in material compliance with MDB's at approval requirements (including implementation of the environmental action program, if any), and (b)(i) mitigation prospects are uncertain or unlikely; or (ii) earlier non-compliance (subsequently corrected) resulted in substantial and permanent environmental damage. No opinion possible: Where, after best efforts, the relevant information to establish material compliance (or lack thereof) cannot be obtained, a rating of no opinion possible may be assigned. This rating should be a last resort, after reasonable effort has been made to obtain the necessary information. A sponsor's failure to report should result in a partly unsatisfactory or unsatisfactory rating only if the sponsor has repeatedly refused to cooperate on this issue.	Financially benign operations, e.g., investments in insurance companies, are not rated. **This standard should be split into a good practice standard that would specify the benchmark for a satisfactory rating and a best practice standard that would specify the benchmarks for excellent, partly satisfactory and unsatisfactory.** **Add separate standard providing benchmarks for financial markets projects, particularly loans to and equity investments in sub-projects.** **Consider adding "not applicable" as an option. See, e.g., IFC standards.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
4.4		*Indicators and benchmarks for MDB's investment profitability:*	
4.4.1	Harmonization	*Good practice.* MDB investment's profitability is based upon the investment's *gross profit contribution* (net of financing costs and loss provisions but before deducting administrative costs) with ratings benchmarks set in relation to corresponding at-approval standards for minimally satisfactory expected performance. *Best practice #1.* If reliable transaction cost data are readily available from management information system, MDB investment's profitability is also based on the investment's *net profit contribution* (the gross profit contribution less administrative costs), measured in risk-adjusted, discounted cash flow terms with ratings benchmarks set in relation to at-approval standards for minimally satisfactory expected performance. *Best practice #2.* If reliable transaction cost data are readily available from management information system, MDB investment's profitability is also based on the investment's *net profit contribution (the gross profit contribution less administrative costs) in relation to the capital employed for the investment,* measured in risk-adjusted, discounted cash flow terms with ratings benchmarks set in relation to at-approval standards for minimally satisfactory expected performance.	In the absence of a standard established by management, I have been willing to accept a proxy standard for minimally satisfactory performance established by the CED. **Standard should be clarified:** • **Meaning of "at-approval standards for minimally satisfactory expected performance."** • **Applicability of "net of financing costs" to equity investments and guarantees.** **See AsDB guidelines for examples of possible proxy standards that might be considered.**
4.4.2	Harmonization	*Good practice.* Loan and guarantee performance benchmarks are set in relation to the MDB's expectations at approval. *Best-practice:* Loan's net profit contribution (net of transaction and financing costs) is sufficient in relation to the MDB's return on capital employed target.	
4.4.3	Harmonization	*Good practice.* Equity investment rates of return are benchmarked (or discounted if an NPV measure is used) against standards for minimally satisfactory expected performance at approval. *Best practice 1.* Equity investment rates of return are benchmarked (or discounted if an NPV measure is used) to reflect appropriate spreads over actual or notional loan yields for the same credit risk, in line with the policy-defined, at-entry approval standard. *Best practice 2.* Where the MDB's investment features both a loan and an equity investment, their combined net profit contribution (net of transaction and loan financing costs) is sufficient in relation to the MDB's return on capital employed target.	Applicable only to institutions that make equity investments. Many institutions have not established policies defining at-entry approval standards for equity investments. A CED facing this problem might, until corporate standards are formally established, seek to determine the minimum general threshold effectively used for equity investments at approval. In the absence of a standard established by

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
			management, I have been willing to accept a proxy standard for minimally satisfactory performance established by the CED. **Too much overlap with GPS 4.4.1.** **Best practice 2 should be in a separate standard. The good practice standard relates to equity investments. Best practice 2 relates to combined loan and equity investments. The GPS should probably also provide a good practice alternative.**
4.5		***Indicators and benchmarks for MDB's additionality:***	
4.5.1	Harmonization	The rating for the MDB's additionality takes into consideration four indicators: • Terms. Would the client have been able to obtain sufficient financing from private sources on appropriate terms? Judgments on this indicator consider pricing (including additional costs arising from MDB conditions that would not be imposed by private investor), tenor, grace period, currency, and timeliness, i.e., the availability of financing without unduly delaying the project. • Was the MDB (because of its being an MDB) needed to reduce the risks or provide comfort (i.e., improve the investors' perceptions of the risks involved) and, thus, to encourage the investors and lenders to proceed? • Was the MDB needed to bring about a fair, efficient allocation of risks and responsibilities, e.g., between the public sector and the private investors? • Did the MDB, improve the venture's design or functioning--in business, developmental, transition, social or environmental terms?	An MC rating requires that additionality be rated separately. Focusing on the underlying objective of fostering comparability among Members, I have not insisted that additionality be covered by XASRs and PERs, but I have required that XASR-As treat additionality as a separate dimension, that additionality not be taken into consideration in the XASR-As' work quality ratings, and that the CED's annual reviews report on the additionality ratings and the work quality ratings shown in the XASR-As (i.e., based on the GPS). The wording of the 4th edition should be clarified to eliminate contention in future benchmarking exercises.
4.5.2	Harmonization	The project's additionality is rated using benchmarks substantially consistent with the following: • Excellent: It is clear that (a) the project would not have gone ahead without the MDB *or* (b) absent the MDB, (i) it would have entailed a materially unfair or inefficient allocation of risks and responsibilities or (ii) it would have been materially weaker in business, developmental, transition, social or environmental terms. • Satisfactory: It is likely that (a) the project would not have gone ahead without the MDB *or* (b) absent the MDB, (i) it would have entailed an unfair or inefficient allocation of risks and responsibilities or (ii) it would have been weaker in business,	An MC rating requires that additionality be rated separately. Focusing on the underlying objective of fostering comparability among Members, I have not insisted that additionality be covered by XASRs and PERs, but I have required that XASR-As treat additionality as a separate dimension, that additionality not be taken into consideration in the XASR-As' work quality ratings, and that the CED's annual reviews report on the additionality ratings and the work quality

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		developmental, transition, social or environmental terms. • Partly unsatisfactory: It is likely that (a) the project would have gone ahead without the MDB *and* (b) the MDB made no contribution to (i) a fair or efficient allocation of risks and responsibilities and (ii) the business, developmental, transition, social or environmental performance of the project. • Unsatisfactory: It is clear that (a) the project would have gone ahead without the MDB *and* (b) the MDB made no contribution to (i) a fair or efficient allocation of risks and responsibilities and (ii) the business, developmental, transition, social or environmental performance of the project.	ratings shown in the XASR-As (i.e., based on the GPS). The wording of the 4th edition should be clarified to eliminate contention in future benchmarking exercises. **The benchmarks for excellent and satisfactory should be revised to incorporate "with financing on appropriate terms" and "without undue delays."** **The standard should be split, with a good practice standard specifying the benchmark for a satisfactory rating and the other benchmarks relegated to a best practice standard.**
4.6		***Indicators for MDB's work quality:***	
4.6.1	Harmonization	The rating for the MDB's work quality is based partly on *at-entry screening, appraisal and structuring work*; i.e., how effectively the MDB carries out its work prior to approval of the investment.	
4.6.2	Harmonization	The rating for the MDB's work quality is based partly on its *monitoring and supervision quality*, i.e., how effectively the MDB carries out its work after approval of the investment.	
4.6.3	Harmonization	The rating for the MDB's work quality is based partly on its *role and contribution*, i.e., the quality of the MDB's contributions from inception to evaluation. The rating judgment considers compliance with basic operating principles, the MDB's contribution to client capacity building objectives (as relevant), the operation's consistency with furtherance of the MDB's corporate, country and sector strategies, and its clients' satisfaction with the MDB's service quality.	
4.6.4	Harmonization	Assessments of the MDB's work quality should be made independently of the ratings assigned for development or transition outcomes and MDB's investment profitability. These assessments, which are benchmarked against corporate good practice, reflect the quality of the MDB's contributions to good or bad outcomes, not the good or bad outcomes themselves.	
4.7		***Standard XASR attachments***	

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
4.7.1	Harmonization	The XASR includes an attachment providing details supporting the financial and economic rate of return estimates (with transparent assumptions and cash flow statements). This attachment provides the basis for review and independent verification of the XASR's judgments and conclusions.	I accepted electronic links to projections. **Should make clear that this standard applies also to PERs.**
4.7.2	Harmonization	The XASR includes an attachment providing, for each safeguard dimension addressed in the MDB's environmental and social guidelines, a comprehensive summary of environmental, worker health and safety, and social outcome compliance information with sufficient evidence from a field visit and/or client reporting to support the assigned outcome and related MDB work quality ratings. This attachment provides the basis for review and independent verification of the XASR's judgments and conclusions.	An electronic link this document would be acceptable. **Should make clear that this standard applies also to PERs.**
5		**Annual Reporting and Process Transparency**	
5.1		***Annual synthesis reporting: Annual Review***	
5.1.1	Harmonization	CED prepares an annual review addressed to the MDB's management, staff and Board of Directors. The scope of the annual review includes, *inter alia*, a synthesis of the CED's validated findings from all XASRs and PERs generated and reviewed during the period covered.	MC rating requires coverage of all "the CED's validated findings from all XASRs and PERs generated and reviewed during the period covered." **Should make clear that "period covered" means the year covered.**
5.1.2	Harmonization	The annual review should provide sufficient information to make the reader aware of possible biases in the sample of projects covered by the annual review. Consequently, the annual review: • Describes how the population was identified and how the sample was selected. If stratification was applied or part of the sampling was non-random, the annual review states the rationale. • Reports on the number of XASRs and PERs for the year. • Includes an annex profiling the important characteristics of the evaluated sample (e.g., sector, investment size, and percentage of operations affected by specific loss provisions) against the population. • Confirms that all projects selected for evaluation had reached early operating maturity, as defined in the GPS, by the time of their evaluation. • If less than 100% of the population has been covered, provides information on statistical confidence levels and states explicitly whether reported success rates can be attributed to the population.	MC requires that practices be consistent with *all* bullet paragraphs. **Penultimate bullet should be revised to call for confirming that all projects that reached early operating maturity have been included in the population in the year in which they reached early operating maturity.** **See also comments on GPS 2.1.2.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
5.1.3	Harmonization	The annual review either (i) describes the ratings criteria and benchmarks in an annex or else (ii) refers to a website providing this information.	
5.1.4	Harmonization	The ratings reported should be those of CED.	**This standard can be dropped. Repeats what is already in GPS 5.1.1. The latter could be modified slightly to ensure that this idea is retained.**
5.1.5	Harmonization	*Good practice.* CEDs should disclose the differences between CED and operating staff ratings and the materiality of the differences. Where CED ratings are reflected partly in XASR-As and partly in PERs, the CED should disclose the differences between CED and operating staff ratings separately for the XASRs and the PERs. The disclosure is made in global terms, not on a project-by-project basis and is limited to differences in binary outcome, additionality and work quality success ratings.	Applicable only to institutions that prepare XASRs, i.e., that do not limit evaluations to PERs. **To eliminate some misunderstandings found in the present benchmarking review, this standard should be more precise. It should (a) explain what is meant by binary, (b) add "development" before "outcome" on the penultimate line, (c) replace "work quality success" with "work quality, and (d) call for disclosure on each dimension separately.**
5.1.6	Harmonization	*Good practice:* For each rating dimension and indicator, the annual review shows the *proportion of the evaluated sample* in each performance-rating category. *Best practice:* The annual review also shows, by dimension, the *proportion of total disbursed MDB financing* for the sample that is in each performance-rating category.	Alternatively, raw numbers can be used by institutions with small number of projects.
5.1.7	Other	*Good practice:* The annual review analyzes the evaluation results and highlights the findings. In doing so, it notes whether findings are statistically significant. *Best practice #1*: The annual review provides a synthesis description of the ratings patterns and their cross-cutting performance drivers under each indicator. *Best practice #2*: The annual review provides the ratings for the previous few years to show how performance is evolving.	**To eliminate some misunderstandings found in the present benchmarking review, should add "dimension and indicator" before "ratings" in Best Practice #2.**
5.1.8	Other	The annual review makes recommendations to Management and the Board based on the evaluation findings.	
5.1.9	Other	CED maintains a tracking system for recording disposition by Management of each	

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
		recommendation.	
5.2		***Process transparency: Annual Report.***	The annual report can be included in the annual review if an MDB wishes to do so.
5.2.1	Harmonization	CED reports periodically (at least every three years) to the MDB's management and Board on the quality and efficacy of the MDB's evaluation system, including the self-evaluation system, any gaps in coverage of the MDB's operations, the work of CED, and the generation and application of lessons learned in new operations. In addition, it submits to the MDB's management and Board the periodic benchmarking reviews of the consistency of the MDB's practices with the GPS.	MC requires that practices be consistent with *all* elements of standard.
6		**Identification of Lessons, Dissemination, and Ensuring Application of Lessons**	
6.1		***Identification of lessons:***	
6.1.1	Other	Lessons should be concise, prescriptive, and placed in the context of a material issue that was encountered in the evaluation so that its relevance to new operations can be determined easily, on a stand-alone basis.	
6.1.2	Other	The point of view and selectivity should focus on what the MDB might have done to obtain better results from the operation.	
6.2		***Dissemination of findings and lessons:***	
6.2.1	Other	*Good practice*: The CED makes available to MDB staff the findings and lessons derived from the MDB's evaluation work. *Best practice*: The CED makes available to MDB staff a range of user-friendly dissemination products covering the XASR and PER findings, the annual review and CED special studies, e.g., access to the full reports, an on-line searchable lessons retrieval network, electronic notification of new items, and PowerPoint slide-shows of annual review or special study findings.	

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
6.3		***Ensuring application of lessons:***	
6.3.1	Other	*Good practice.* It is the responsibility of operational department managers to ensure that past lessons have been systematically researched, identified and applied in new operations. *Best practice #1*. Standard processing documentation for new operations includes a prompt, *in early stage documents*, for relevant past lessons, complemented by a prompt, in final decision-stage documentation, for how the past lessons have been addressed in the appraisal and structuring of the new operation. *Best practice #2:* Procedures call for CED to review documents on new operations with respect to identification of relevant lessons from evaluated operations.	
6.3.2	Other	In its annual evaluation process report, CED reviews and reports to management and the Board the evidence available for judging the extent to which lessons are being incorporated in new operations.	
6.3.3	Other	*Good practice.* Internal corporate reporting (up to Board level) is broadly aligned with the evaluative framework. *Best practice #1:* Reports, from project-level to department- and corporate-level, cover development or transition outcome, investment outcome, additionality and MDB work quality. *Best practice #2*. Reports apply coherent and consistent benchmarks across projects and at all stages of the project cycle (appraisal, supervision/monitoring and evaluation).	Similar to GPS 1.1.3, but GPS 1.1.3 relates to input by CED, and GPS 5.4.3 relates to outcomes and internal coherence. This integration of evaluative scope, measurement standards, findings and reporting with corporate- and unit-level portfolio reporting caters for the results-based management principle and reality that "what gets measured, gets done," and that properly "what gets done, gets measured coherently and consistently." Without this integration, a disconnect is likely between predominant operational and career incentives and application of evaluation lessons for getting better outcomes. **This standard does not belong under Ensuring Application of Lessons.** **The good practice standard can be dropped and the two best practice standards converted to good practice standards. Since they are "other" standards, they will not come into the scoring in the periodic benchmarking reviews.**

CRITERIA USED FOR APPLYING GOOD PRACTICE HARMONIZATION STANDARDS AND SUGGESTIONS FOR POSSIBLE CHANGES IN STANDARDS

Std. #	Type of Std.	Standard	Criteria Used for Applying Standards & (in bold font) Suggestions for Possible Changes in Standards
6.4		***Disclosure***	
6.4.1	Other	To protect client company confidentiality, the candor needed for effective corporate learning, and the risk to the MDB's credit rating that partial release of investment portfolio data (and related standards and benchmarks) might entail, the MDB does not disclose individual evaluation reports or the full text of the CED's annual review.	
6.4.2	Other	*Good practice:* The MDB's disclosure policy for evaluation products should be explicit, should be consistent with the MDB's general disclosure policy, and should cover all evaluation products. *Best practice*: The MDB's disclosure policy is disclosed via the CED's web page, specifically noting any exceptions applicable to evaluation reports.	
6.4.3	Other	*Good practice:* The MDB includes an accurate summary of CED's major annual review findings in its Annual Report. *Best practice 1:* CED prepares and posts on the MDB's external website an abstract of its annual review that accurately summarizes its essential findings, including the outcome, additionality and work quality ratings profiles, sampling representativeness, ratings criteria, benchmarks, and consistency with core GPS. Best practice 2: After appropriate redaction to protect commercial confidentiality, CED discloses its aggregate evaluation reports.	This standard seeks to ensure that the CED's evaluation findings are readily available to the public. The good practice standard calls for the MDB's annual report to include an accurate summary of the CED's findings. The second "best practice standard" calls for posting on the MDB's external website an abstract of the CED's annual review that accurately summarizes its essential findings. Since the best practice standard serves the same purpose as the good practice standard, I have been willing to assign a MC rating for this standard provided that the MDB's annual report refers to the website for access to the annual review and provided that the information posted meets the conditions of the best practice standard. **This interpretation should be incorporated into the GPS.**

COMPARABILITY OF RATING CATEGORIES

Std. #	Standard	Shortfalls in Comparability
4.3.2- 4.3.17	***The development or transition outcome*** is based on: • The project's contribution to the company's business success; • The project's contributions (positive or negative) to the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy; • The project's contributions to economic development; • The company's overall environmental impacts in the area of influence, or the vicinity, of the project at the time of evaluation.	**AfDB**. Comparability limited. Three of the four indicators are not comparable, but their lack of comparability is partly overcome in the synthesis rating for the development outcome because un-synthesized elements relating to individual indicators are synthesized at the development outcome level. See comments on GPS 4.3.2, 4.3.8, 4.3.10 & 4.3.16, below. **AsDB**. Comparability limited. See comments on GPS 4.3.8 & 4.3.16, below. **EBRD**. Not comparable because components of EBRD's "transition outcome" are not comparable with components of GPS (GPS 4.3.2, 4.3.8, 4.3.10 & 4.3.16, below). In particular, EBRD does not rate contributions to the company's business success and economic contributions. (EBRD reports primarily on "transition impact," rather than "transition outcome," but also provides information on "transition outcome," which is closer to the GPS category, in an annex to its annual review. The two concepts differ because (a) "transition impact" includes additionality as an indicator of the "results on the ground," but "transition outcome" treats additionality as a separate evaluation dimension, and (b) "transition impact" excludes environmental impacts, but "transition outcome" includes them.) **MIGA**. Comparability limited. See comments on GPS 4.3.2 & 4.3.16, below.
4.3.2	The project's development or transition outcome is based partly on the ***project's contribution to the company's business success***, measured primarily by the real after tax returns and secondarily by (i) the project's contribution to other business goals articulated at approval and (ii) the project *company's* prospects for sustainability and growth.	**AfDB.** Not comparable because AfDB (a) rates business success as a separate dimension, rather than as one of the indicators for assessing development outcome, (b) does not rate business success "primarily by the real after tax returns," but rather suggests several financial ratios for assessing business success, and (c) does not consider "the project company's prospects for sustainability and growth." **EBRD.** The project's contribution to the company's business success is not relevant for EBRD because its Evaluation Department has stated that EBRD does not seek to contribute to economic development and poverty reduction. It does, however, rate project and company financial performance based on (a) the financial internal rate of return, (b) the economic internal rate of return, and (c) the project's sales, profitability, solvency, and debt/equity ratio. It also rates "fulfillment of project objectives" as a separate indicator but does not rate the company's prospects for sustainability and growth. The indicators for EBRD's "project and company financial performance" (particularly items (b) and (c), above) are not comparable with those in the GPS.

COMPARABILITY OF RATING CATEGORIES

Std. #	Standard	Shortfalls in Comparability
		MIGA. Comparability limited because MIGA's guidelines call for considering the project's contribution to other business goals articulated at approval only when no quantitative information is available and do not call for considering the project company's prospects for sustainability and growth.
4.3.8	The project's development or transition outcome is based partly on the ***project's contributions (positive or negative) to the country's private sector development and/or its development of efficient capital markets and/or its transition to a market economy.***	**AfDB**. Not comparable because AfDB (a) rates contributions to private sector development and contributions to the enabling environment separately and does not call for an overall rating and (b) does not call for consideration of negative as well as positive outcomes. **AsDB**. Comparability limited because (a) AsDB's list of indicators differs from GPS list and (b) its standard does not clearly call for consideration of negative outcomes. **EBRD**. Although EBRD uses most of the indicators specified in the GPS, comparability is still limited because it (a) does not specifically call for considering contributions to transition resulting from "development of financial institutions and financial markets," which it considers to be subsumed in other indicators, and (b) does not consider contributions to transition resulting from "development of physical infrastructure used by other private parties," because "development of physical infrastructure is not an objective for EBRD."
4.3.10	The project's development outcome is based partly on **its *contributions to economic developmen*t**. Performance is assessed not only by the direct economic costs and benefits to the owners and financiers but also by the economic costs and benefits to customers, employees, government, suppliers, competitors, local residents, etc.	**AfDB.** Not comparable because AfDB rates growth in the economy & living standards separately and does not call for an overall rating. **EBRD.** The project's contributions to economic development are not relevant for EBRD because its Evaluation Department has stated that EBRD does not seek to contribute to economic development and poverty reduction. It some cases, it considers the economic internal rate of return as an indicator in rating project and company financial performance, but it does not show separate ratings for this indicator.

COMPARABILITY OF RATING CATEGORIES

Std. #	Standard	Shortfalls in Comparability
4.3.16	The project's development or transition outcome is based partly on the company's overall ***environmental performance*** at the time of evaluation. The assessment is based primarily on the MDB's specified standards in effect at investment approval (though compliance with the standards specified at the time of the evaluation may be taken into consideration in assigning "outstanding" or "excellent" ratings). "Environment" includes the physical environment and, to the extent covered by the MDB's policies, also includes energy efficiency; the quality of the client's environmental management activities; social, cultural, health and safety impacts; and the extent of public consultation and participation.	**AfDB**. Comparability limited because AfDB calls only for assessing the environmental outcome of project, i.e., it ignores the broader issue of environmental impacts in the area of influence or the vicinity of the project. **AsDB**. Comparability limited because AsDB (a) calls only for assessing the environmental outcome of the project, i.e., it ignores the broader issue of environmental impacts in the area of influence or the vicinity of the project, and (b) includes poverty reduction here, rather than as a part of contributions to economic development. **EBRD**. Comparability limited because EBRD takes into consideration here as well as under work quality its role with respect to environmental and social matters. **MIGA**. Comparability limited because MIGA assesses performance against its at-approval and at-evaluation requirements for all ratings, rather than just for its excellent ratings.
4.4.1	**The *MDB investment's profitability*** (contribution to its corporate profitability objective). *Good practice.* MDB investment's profitability is based upon the investment's *gross profit contribution* (net of financing costs and loss provisions but before deducting administrative costs) with ratings benchmarks set in relation to corresponding at-approval standards for minimally satisfactory expected performance.	**AsDB**. Not comparable because AsDB's ratings can use 3 alternative proxies for "at-approval standards." Not all can be equivalent to "at approval standards for minimally satisfactory expected performance." **EBRD**. Not comparable because EBRD does not base its rating on "at-approval standards for minimally satisfactory expected performance." **MIGA** has not issued final guidelines for this item. Moreover, its draft guidelines relate to Best Practice Standard No. 1, rather than to the Good Practice Standard. At this point, it has nothing comparable with the Good Practice Standard.
4.5.1	**The *MDB's additionality.*** The rating for the MDB's additionality takes into consideration four indicators: • Terms. Would the client have been able to obtain sufficient financing from private sources on appropriate terms? Judgments on this indicator consider pricing (including additional costs arising from MDB conditions that would not be imposed by private investor), tenor, grace period, currency, and timeliness, i.e., the availability of financing without unduly delaying the project. • Was the MDB (because of its being an MDB) needed to reduce the risks or provide comfort (i.e., improve the investors' perceptions of the risks	**AfDB.** Not comparable because AfDB treats additionality as part of work quality, rather than as a separate dimension. **AsDB**. Comparability limited because AsDB doesn't consider whether (a) alternative financing was available on appropriate terms and (b) whether AsDB was needed to bring about a fair, efficient allocation of risks & responsibilities. **EBRD.** Limited comparability because benchmarks differ from GPS. **MIF**. Comparability limited because MIF's guidelines do not consider whether MIF was needed (a) to reduce the risks or provide comfort and (b) to bring about a

COMPARABILITY OF RATING CATEGORIES

Std. #	Standard	Shortfalls in Comparability
	• involved) and, thus, to encourage the investors and lenders to proceed? • Was the MDB needed to bring about a fair, efficient allocation of risks and responsibilities, e.g., between the public sector and the private investors? Did the MDB, improve the venture's design or functioning--in business, developmental, transition, social or environmental terms?	fair, efficient allocation of risks and responsibilities. **IFC**. Not comparable because IFC treats additionality as part of work quality, rather than as a separate dimension. **MIGA**. Not comparable because MIGA treats additionality as part of work quality, rather than as a separate dimension.
4.6.1	The rating for the MDB's work quality is based partly on ***at-entry screening, appraisal and structuring work***; i.e., how effectively the MDB carries out its work prior to approval of the investment.	**MIGA.** Comparability limited because MIGA assigns a single rating with respect to both GPS 4.6.1 and 4.6.2.
4.6.2	The rating for the MDB's work quality is based partly on its ***monitoring and supervision quality***, i.e., how effectively the MDB carries out its work after approval of the investment.	**MIGA.** Comparability limited because MIGA assigns a single rating with respect to both GPS 4.6.1 and 4.6.2.
4.6.3	The rating for the MDB's work quality is based partly on its ***role and contribution***, i.e., the quality of the MDB's contributions from inception to evaluation. The rating judgment considers compliance with basic operating principles, the MDB's contribution to client capacity building objectives (as relevant), the operation's consistency with furtherance of the MDB's corporate, country and sector strategies, and its clients' satisfaction with the MDB's service quality.	**AfDB.** Not comparable because AfDB treats additionality as part of work quality-role and contribution, rather than as a separate dimension. **AsDB**. Comparability limited because AsDB does not consider the operation's consistency with furtherance of the its corporate, country and sector strategies. **EBRD**. Comparability limited because EBRD does not consider (a) compliance with its basic operating principles, (b) consistency with furtherance of its corporate, country and sector strategies, and (c) its clients' satisfaction with its service quality. **IFC**. Not comparable because IFC treats additionality as part of work quality-role and contribution, rather than as a separate dimension. **MIGA**. Not comparable because MIGA treats additionality as part of work quality-role and contribution, rather than as a separate dimension.