

**Good Practice Standards for Evaluation of MDB Supported Public Sector Operations**

**ADDENDUM**

**Good Practices for the Evaluation of Policy-Based Lending**

**Multilateral Development Bank Evaluation Cooperation Group**

**December 2004**

## ABBREVIATIONS

AfDB	-	African Development Bank
AsDB	-	Asian Development Bank
ECG	-	Evaluation Cooperation Group
GPS	-	good practice standards
IMF	-	International Monetary Fund
MDB	-	multilateral development bank
PBL	-	policy-based lending
PBLs	-	policy-based lending operations
TA	-	technical assistance

## CONTENTS

A.	Introduction	1
B.	Evaluating Policy-Based Lending Operations	2
	1. Why policy-based lending operations merit special interpretation	2
	2. Good Practice Standards and Treatment of Policy-Based Lending Operations	2
C.	Evaluation of Individual Operations	3
	1. Timing	3
	2. Evaluability	3
	3. Assessing Inputs and Design Processes	4
	4. Capturing Ownership	5
	5. Political Economy of Decision Making	5
	6. Assessing Outputs	5
	7. Assessing Outcomes	6
	8. Assessing Impacts	6
	9. Evaluating Policy-Based Lending Process	7
D.	Rating Policy-Based Lending Operations	
	1. Criteria and Subcriteria	8
	2. Criteria, Aggregation, and Weighting	9
	3. Minimum Thresholds	10
E.	Management of Policy-Based Lending Evaluations	10
	1. Coverage	10
	2. Fostering Usage and Accountability	11
	3. Disclosure	11
APPENDIX		
	Subcriteria for Evaluating Policy-Based Lending Operations	12

## A. INTRODUCTION

1. The March 1996 report of a Task Force on Multilateral Development Banks noted that it was not possible to compare or even describe operational results across the multilateral development banks (MDBs).<sup>1</sup> It was recommended that common methodologies for evaluating MDB portfolios be adopted, building on best evaluation practices. In response to this recommendation, the Evaluation Cooperation Group (ECG) was formed; it initially consisted of the heads of the evaluation offices of the five MDBs referred to in the Task Force Report: the African Development Bank (AfDB), the Asian Development Bank (AsDB), the European Bank for Reconstruction and Development, the Inter-American Development Bank, and the World Bank. The European Investment Bank joined in 1998, and the International Monetary Fund (IMF) in 2001. In the spirit of fostering harmonization and adoption of good practices, the ECG undertook an examination of the evaluation of MDB-supported public sector operations, resulting in 2002 in a report on good practice standards (GPS).<sup>2</sup> The GPS report recognized that policy-based lending (PBL) was a special category of public sector lending, and recommended that special guidelines be prepared to “*meet the special evaluation requirements for such operations, if necessary (p.13).*”

2. In 2003, the ECG agreed that a study would be carried out to take stock of evaluation practices for PBL in the MDBs, to assess the evaluation approaches being used and the management of the evaluation process for PBL. The resulting study, managed by the AsDB, was finalized in September 2004.<sup>3</sup> This Addendum to the GPS provides good practices for the evaluation of PBL, including the rating of individual policy-based lending operations (PBLs) and the management of the evaluation process for PBLs.

3. PBL has evolved over the last two decades. Its objectives have evolved from stabilization to encompass support for institutional development and changes in development policy. This has resulted in operations that are more complex in both design and implementation. Policy and institutional reform, as now supported under PBL, is an ongoing process, sensitive to factors such as ownership, political economy of decision making, participation, and the quality of PBL design. A number of lending modalities have been developed as variations on the basic purpose of facilitating policy change, including crisis modalities and hybrid modalities that include an investment component. PBL plays a larger role than ever before in MDB operations, even than during the crisis periods of the late 1990s. For the regional development banks with such modalities, PBL is now more than 20% of approvals by loan amount; for the World Bank it is now more than 30% of approvals.

4. While MDB evaluations of PBL differ, the conceptual framework is similar. Assessment of PBL performance, from the various levels of inputs to impacts, is used to provide accountability and to distill lessons. Standard criteria, such as relevance, efficacy, efficiency, sustainability, institutional development, impact, borrower performance, and MDB performance, are applied to assess overall performance, and, in some MDBs, to rate PBLs. However, there is no single set of agreed upon criteria, methods, and techniques for the evaluation of PBL. Although the same general criteria may be applied to different forms of investment project or

---

<sup>1</sup> Task Force of the Development Assistance Committee of the Organization for Economic Cooperation and Development.

<sup>2</sup> Evaluation Cooperation Group. 2002. *Good Practice Standards for Evaluation of MDB-Supported Public Sector Operations*.

<sup>3</sup> Steven Tabor and Stephen Curry. 2004. *Good Practices for the Evaluation of Policy-Based Lending by Multilateral Development Banks*. Prepared for the ECG. Manila. AsDB.

PBL evaluation, the way in which this is done allows for considerable variability in the assessment of PBL.

## **B. Evaluating Policy-Based Lending Operations**

### **1. Why Policy-Based Lending Operations Merit Special Interpretation**

5. From an evaluation perspective, PBLs differ from investment operations in many respects. Disbursement in PBLs is against completion of reform measures, compared with disbursement under investment projects, where payments are usually made directly to contractors and suppliers against authorized claims for work already done or supplies already delivered. PBLs are complex operations that impact upon a large number and array of stakeholders. PBLs may be motivated by a need to contribute to the restoration of macroeconomic and political stability, although this may not be explicitly declared in the project documentation. It is more difficult to assess, isolate, and attribute the impacts of PBL, because PBLs aim to make non-marginal changes to the rules and incentive systems under which economies operate. Rarely is there a “control” against which progress can be measured. Policy and institutional reforms generally take longer than the life of any PBL operation to implement or for the impacts to become evident.

6. PBL preparation and implementation are often compressed into a shorter time period than for investment operations. The combination of short preparation periods and time-sensitive implementation limits the extent to which PBL design may include baseline studies, ex-ante impact assessment, or other inputs that provide a solid foundation for self- and independent evaluation. Moreover, the theory underlying the dynamic linkages between complex, multifaceted policy and institutional reform packages and socioeconomic performance is not well established.

### **2. Good Practice Standards and Treatment of Policy-Based Lending Operations**

7. The GPS established a set of criteria and procedures for the evaluation of both individual and broader public sector operations. These criteria and procedures are more straightforward to apply in the evaluation of investment projects, where there are clear project-level objectives, implementation schedules, and procurement processes, than in the evaluation of PBL, where objectives are broader, implementation schedules less precise, and procurement processes undefined. Given the special purposes of PBL, and its specific characteristics, the GPS need interpretation and modification for the evaluation of PBL operations.

8. A PBL operation is assessed as an external assistance instrument in which agreed upon inputs contribute to outputs, which in turn give rise to outcomes, which in turn contribute to impacts. PBL inputs include agreed upon policy conditions, finance, and technical assistance (TA) where applicable. PBL outputs refer to the implementation of reforms and the disbursement and utilization of PBL finance. PBL outcomes refer to changes in the policy or institutional enabling environment that occur as a result of the implementation of agreed upon reforms. This could include, for example, changes in the role of the public and private sectors, governance practice, competitiveness, openness, public expenditure incidence, prudential soundness of the financial system, stakeholder voice in decision making, sector development regimes, and/or distribution of decision-making authorities between central and local governments. PBL impacts refer to changes in economic, social, environmental, and human development performance directly attributable to PBL-supported policy reform. Changes in

economic output, incomes, exports, inflation, health and education levels, poverty incidence, and environmental conditions are some of the types of impact variables that are assessed in PBL evaluations. Assessment of PBL performance, from the various levels of inputs to impacts, is used to draw conclusions about, and distill lessons from, the performance of an operation.

## **C. Evaluation of Individual Operations**

### **1. Timing**

9. PBLs generate outcomes that become apparent only some time after a program is completed. Impacts may take even longer to be realized. Sufficient time must have elapsed since completion of a PBL operation for outcomes and impacts to be realized and for the sustainability of the operation to be apparent. If PBL evaluations are conducted too soon, there may be inadequate time for reactions to the policy change. If evaluations are conducted too long after completion, it may be difficult to isolate the effects of the PBL operation. PBLs should be self-evaluated shortly after completion and postevaluated some 2-3 years after completion because

- evaluation undertaken shortly after completion is adequate for assessing whether or not reforms are under way—i.e., if inputs have been delivered and initial outputs have resulted; and
- 2-3 years after completion is generally sufficient to assess outputs, outcomes, impacts, and the sustainability of PBL-supported reforms.

10. Where there is a series of overlapping and related PBLs, the timing question is more complicated, since the policy reform effort will have been supported by several operations. In such instances, a PBL operation, and the series of PBLs, can be evaluated individually or as part of a cluster of related operations, depending on a judgment of when the outcomes and impacts of the sequence of related PBLs are likely to be realized. It is recommended that

- cluster evaluations of individual operations within a series of programmatic PBLs should be timed to ensure that sufficient time has lapsed for outcomes and impacts to be realized;
- clusters of related PBLs should be evaluated in sufficient time to be an input in the preparation of country assistance operations; and
- more frequent independent evaluations of related PBLs are warranted in instances in which completion reports highlight substantial difficulties or slippages in implementing the agreed upon strategy or in instances in which the PBLs provide particularly valuable lessons for replication (or avoidance) in other countries.

### **2. Evaluability**

11. PBL evaluability can be improved by taking steps to reveal key unwritten objectives, and to include these in the set of objectives against which the PBL will be evaluated. Key PBL objectives may be both unwritten and ambiguous. Implicit motivating factors can include, for example, augmenting liquidity provided by a lead financier, staving off a debt default, providing comfort to other agencies and the private sector that economic management is on track, and ensuring that a fiscally distressed government has sufficient resources to meet the counterpart financing requirements of other ongoing MDB operations.

12. For many development policy-oriented PBLs, a key implicit objective is to provide resources to governments to improve the balance of payments, to contribute to political stability, or to enhance fiscal performance. While macroeconomic objectives are fully explicit in the case of IMF stabilization operations, they tend to be either implicit in development policy-oriented PBLs or subsumed under a broadly defined goal of “restoring sound macroeconomic performance.” Implicit macroeconomic objectives are always important, and should be addressed. It is recommended that a PBL should be evaluated with respect to its contribution to improved macroeconomic balances, whether this is stated in the objectives or not. This would include an assessment of the reasonableness of the macroeconomic assumptions and program, and the performance of the authorities in correcting macroeconomic imbalances. It would also include any unintended as well as intended outcomes and impacts.

13. Often PBLs do not include logical or results frameworks or clear measures of policy reform results. MDB evaluators should encourage operations staff to specify and justify the expected results of their operations. In those cases where a PBL framework was absent or poorly defined, it is recommended that

- a retrospective logical or results framework should be prepared for guidance for post evaluation; and
- PBL postevaluations should review the extent to which the PBL made adequate arrangements for monitoring and evaluation, as a part of a broader assessment of the evaluability of the PBL.

### **3. Assessing Inputs and Design Processes**

14. The evaluation of PBL inputs should assess the internal logic of a PBL operation, its consistency with other operations, the validity of underlying assumptions, whether it addressed the relevant development constraints, and whether it was the best option. The design of a PBL operation, in terms of its inputs and the process by which those are derived, can be compared to processes and practices suitable for replication and/or avoidance. Some of the most important design factors influencing high-performance PBLs include the following:

- The national development agenda underpins reform.
- There is ownership and favorable political economy of decision making.
- There is adequate country knowledge and analysis.
- The institutional capacity to manage the reform process and to implement agreed upon reforms is analyzed and factored into the design.
- The program logic includes an identification of the goals to be achieved; the specific purpose of the operation; reform measures; their expected outputs, outcomes, and development results; together with the key assumptions and risks to performance.
- Appropriate reform instruments are selected and properly sequenced to reach policy targets; reform instruments should be internally consistent, detailed, manageable in number, and aimed at substantial reductions in major distortions and early success in delivering results.
- The economic, social, institutional, and environmental impacts are analyzed, and mitigation measures are incorporated where necessary.
- Performance risks are clearly identified, and risk management strategies are agreed upon.
- Adequate external financing is provided to encourage the country to implement reforms with vigor.

- The financing rationale is established, and impacts (debt sustainability, efficiency of marginal public spending, impact on reform pace) are assessed.
- Reporting, monitoring, and evaluation indicators are identified, and related responsibilities are assigned and funded.

PBL design should be treated as part of a long-term reform process that is best reinforced by implementation practices in which the design makes provisions such that (i) supervision can be used as an opportunity for learning and sharing lessons, (ii) participation is fostered, (iii) problems identified during implementation can be expeditiously assessed and resolved, and (iv) the reform agenda builds a constituency for the next generation of reforms.

#### **4. Capturing Ownership**

15. Ownership is the extent to which governments are fully committed to, directly involved in, and accountable for a program of policy reforms. A part of this is the degree to which executing agencies provide consistent support for reform programs when loan resources flow not to them but to the general budget. Ownership is now well recognized as an important determinant of PBL performance. A more systematic and robust set of tools for assessing credible commitment to reform, and for relating this to the design and implementation of PBLs, should be introduced in PBL evaluations. Several methodological approaches, such as leadership analysis, stakeholder analysis, institutional assessment, and reform readiness analysis, can be used for this purpose. Even when formal ownership assessment tools are not applied, PBL evaluations should assess

- the adequacy of the political support accorded to national reform champions;
- the degree of support provided to reforms by executing agencies and agencies responsible for implementation;
- the degree to which the purpose, objectives, and likely effects of reforms were adequately communicated to the public; and
- the extent to which PBL design reinforced national ownership.

#### **5. Political Economy of Decision Making**

16. Conducive political conditions are an important determinant of effective PBL operations. A relatively small set of political economy factors in a borrower country appears to explain the success or failure of a PBL operation. This includes commitment to reform shared by influential people at the political and bureaucratic levels, social cohesion, timing, accountability, length of tenure of the government, and whether or not a crisis precipitated the reform. It is good practice in a PBL evaluation to examine the degree to which

- the political economy of decision making, at design and during implementation, was conducive to effective reform, through assessing the process of policy decision making, the role of reform champions, the likely winners and losers in the reform process, and the nature of the institutional incentives facing those affected by reform.

#### **6. Assessing Outputs**

17. PBL evaluations should include a thorough evaluation of the extent to which PBL outputs were actually achieved. In assessing PBL outputs, it is recommended that

- evaluators should assess not only the extent to which inputs were delivered (i.e., agreed upon reforms took place), but also the degree to which complementary measures necessary for their implementation occurred; this may include a number of complementary changes in legislation, regulation, public awareness, and institutional arrangements.

## **7. Assessing Outcomes**

18. PBL outcomes can be identified and assessed by reviews of performance indicators, activity surveys, and structured interviews. In addition to assessing the degree to which intended outcomes did transpire, it is recommended that

- observed outcomes should be benchmarked against regional or international standards of public policy and institutional performance to assess the significance of PBL-supported reform to transforming policy settings.

19. Changes in public policy can have wide-ranging and lagged effects, with both desired and perverse outcomes. Many factors besides the PBL operation may influence outcomes. Shocks, macroeconomic developments, other policy reforms, and changing market conditions can influence the same set of performance variables as PBLs. PBL performance evaluations should acknowledge the difficulty that this poses for isolating and attributing outcomes to any given PBL operation and should carefully distinguish the effects that can be attributed to PBLs from other factors.

20. Comparing the situation before a PBL operation to that afterwards has emerged as the most common technique for evaluating outcomes. The advantage of this approach is that it compares actual results in both situations, but it suffers from the strong assumption that the PBL operation alone was responsible for the observed socioeconomic changes.

## **8. Assessing Impacts**

21. A comparison of what would have resulted without the policy changes to an appropriate counterfactual is the appropriate way to evaluate outcomes and impacts. Individuals intimately involved in a reform process can often identify the counterfactual. It is recommended that

- PBL performance evaluations should include an identification of the probable counterfactual to the reforms and financial (and other) support provided under an MDB-assisted PBL operation.

22. Where factors other than PBL implementation may have influenced socioeconomic effects, it is good evaluation practice to isolate these from the effects of policy change. A variety of alternative techniques can be applied to isolate policy effects from other factors that may have influenced socioeconomic activity. Some PBL evaluations have employed quantitative decomposition methods to isolate the effects of policy change from major shocks and changes in the terms of trade. Others have used cross-country regression models to distinguish the effects of policy change from varied starting points and structural characteristics of borrowers. Regardless of the technique used to isolate other effects from policy change, it is recommended that

- an evaluative judgment must be reached about the degree of importance that policy change, vis-à-vis other factors, has had on socioeconomic performance.

23. While the counterfactual can be used to describe what might have occurred in the absence of PBL, many stakeholders, other than those who are party to a PBL operation, are involved in the policy change process. Increasingly, the notion of formal attribution has been replaced with an assessment of the “incentive value” or “additionality” that PBL has brought to the policy reform process. It is recommended that the additionality of a PBL operation be evaluated in terms of whether or not it

- accelerated (or delayed) reform,
- strengthened the hand and credibility of reformers,
- raised the perceived political returns to reform (in terms of easing budget constraints and positive reputation effects),
- fostered policy learning,
- built domestic capacity to design policy, and
- spurred debate and dialogue on new approaches to meeting development objectives.

24. A range of quantitative and qualitative techniques can be used to assess PBL impact. In many cases, a simple comparison of the actual results with those likely to have resulted if the PBL was not undertaken at all is sufficient to illustrate the order of magnitude of PBL impacts. Quantitative techniques used for assessing impacts include partial equilibrium analysis, representative household models, simulation models, and applied general equilibrium models. Common qualitative techniques include participant interviews and beneficiary satisfaction surveys with the results of policy change “scored” directly by stakeholders. It is recommended that

- a combination of quantitative and qualitative approaches should be used to assess PBL impacts.

25. It is important to assess whether PBLs were sufficiently complete to generate desired higher level impacts (i.e., growth, employment, and private sector development). In some cases, a qualitative assessment of the linkages between the PBL operation and the desired development objectives is sufficient to identify what elements were missing, or could have been better designed. With adequate benchmarks and ex-post performance information, simulations, cost-benefit, cost-effectiveness, and other quantitative techniques can also be used to inform such judgments.

26. Evaluation of PBL poverty impacts draws on a variety of techniques and survey instruments to assess changes in living standards, livelihoods, benefit incidence, service provision, and the performance of safeguard measures. To provide appropriate feedback to government and development partners on poverty reduction efforts

- PBL evaluations need to assess whether or not a reform operation could have been more pro-poor in its design and implementation.

## **9. Evaluating the Policy-Based Lending Process**

27. Transaction costs of a PBL operation, both those incurred by the MDB and those by the government, should be evaluated. Resources released from higher cost lending may be

deployed more effectively in other development support activities. From the borrower's perspective, lower transaction costs reduce the burden on overstretched public administrations.

28. Timeliness and adequacy of disbursements should be evaluated. PBLs are planned in part to contribute to the resolution of balance-of-payments and fiscal difficulties. The extent to which this objective is actually achieved is a function of the timeliness of disbursements. Comparing the actual with the planned disbursement profile, and comparing actual disbursements with measures of fiscal or balance of payments requirements, can be useful in assessing the timeliness of the PBL to closing the resource gap.

29. Since some PBLs earmark resources to defray budget or adjustment costs, the PBL evaluation should assess

- the extent to which adjustment costs did materialize, and
- if PBL-provided resources were sufficient to meet these obligations.

30. The extent to which a PBL, and its associated TA, contributed to fostering institutional development should be covered in performance evaluations. The contribution of a PBL to institutional development can be evaluated in terms of whether or not improved governance practices, or improved skills, procedures, incentives, structures, or institutional mechanisms, came into effect. This should include evaluating the contribution made by the PBL to building the capacity to lead and manage the policy reform process.

31. The MDBs are not the only institutions that evaluate PBLs. PBLs are evaluated by research institutes, government agencies, bilateral agencies, the press, nongovernment organizations, and others. These may generate findings different from those reached by MDB evaluators. Capturing and reflecting the assessments of those affected can serve to broaden the extent to which the voices of different stakeholders are reflected in the evaluation process. It can also help to provide important lessons to the government and the MDB as to why a policy change operation may be perceived differently by other stakeholders. It is recommended that

- evaluations should reflect PBL assessments of those directly affected by the reform process, or with an interest in the results.

#### **D. Rating Policy-Based Lending Operations**

32. Currently, the AfDB, the AsDB, and the World Bank rate individual PBLs at the completion report and performance report stages. Until recently, the AfDB and the AsDB rated all their PBLs at both stages. The World Bank does performance reports on 25% of PBLs. There are differences in the success rates of PBLs in these institutions, substantial enough to indicate different uses of the same criteria. Overoptimistic rating is a concern, especially in self-evaluation subject to limited independent attestation.

##### **1. Criteria and Subcriteria**

33. PBLs can be assessed and rated according to six criteria: relevance, efficacy, efficiency, sustainability, institutional development, and impact. In addition to these six, it is useful to assess PBLs with respect to two additional criteria: the performance of the MDB and the performance of the borrower.

34. Most PBLs are rated as being highly relevant, largely because they typically focus on policy or institutional problems of national importance. Relevance, however, should be assessed in terms of the importance of the policy problems, the appropriateness of the choice of PBL (versus TA or dialogue) to the resolution of those problems, and the degree to which the design of the PBL was relevant to the resolution of the identified policy and institutional deficiencies. If an operation is not found to be relevant in its design, timing, or selection of components, then less importance should be attached to efficacy, efficiency, and sustainability status. PBLs assessed to be of limited relevance are unlikely to receive a high overall rating.

35. In some instances, it may not be possible to obtain an accurate estimate of the efficiency of a PBL operation, at least in comparing all costs and returns of a policy change operation. In practice, efficiency generally can be assessed only in a second-best manner, through an analysis of timeliness of finance, with specific evidence on transaction costs or with partial assessment of reform costs and returns. To the extent that available data permit, efficiency of PBLs, or of major reforms, should be assessed. Where a direct assessment of efficiency is not possible, evaluations should at a minimum assess the efficiency of the preparation and implementation processes.

36. PBLs are designed to have major impacts, either at a sector or an economy-wide level. A PBL operation may have a large positive or negative impact (or institutional effect) whose economic and social importance overrides all other factors under consideration. Moreover, it is the impact of a PBL that is the main measure by which national stakeholders assess the performance of this type of assistance operation. Both the actual and expected impact of a PBL operation should be included in an assessment.

37. Applying the same subcriteria for investment operations and for PBLs is problematic because of differences in the nature of the assistance instrument and in determinants of performance. Defining subcriteria in a way that is applicable to PBLs may help to provide an evaluative framework for more uniform, systematic, and comparable assessment. MDB evaluators have drawn on good practice in policy research and field experience to evolve a set of evaluative subcriteria suitable for assessing PBLs. A list of PBL-specific subcriteria for each of the eight criteria indicated in para. 33 is provided in the Appendix. This list is not meant to be either exhaustive or minimal—it reflects many of the factors found to be important determinants of performance, a subset of which are likely to be suitable in varied settings. An evaluative judgment is required to assess the degree to which chosen subcriteria have been achieved in a particular evaluation.

## **2. Criteria, Aggregation, and Weighting**

38. The GPS report recommends that four criteria be treated as “core” criteria in devising an aggregate rating: relevance, efficacy, efficiency, and sustainability. It recommends that institutional development and impact be treated as additional noncore criteria. Because PBLs have major effects of institutional development, and because the impacts of PBL may override all other factors under consideration, it is here recommended that

- institutional development and impact should be treated as “core” criteria for PBLs also.

This implies that it is preferable to rate PBLs by aggregating over six criteria, rather than just the four core identified in the GPS report.

39. Comparability across PBLs can be enhanced if operations make use of similar criteria. Given that some criteria are more readily assessed in some PBLs than in others, it is recommended that the assignment of weights be left to the evaluator, and an explanation for the weights used to derive an aggregate rating be provided at the time the issues to be evaluated are identified. For those PBLs where efficiency can be measured only at a high cost, the weight for this criterion could be reduced. In all cases, the relative importance of all effects should be made in a fully transparent manner, with a clear justification for the final judgment.

### **3. Minimum Thresholds**

40. It is recommended that rating systems specify minimum performance thresholds for the core criteria for any PBL operation to be rated as partly or fully satisfactory. Failure to achieve a minimum assessment under any of the six core criteria implies that a PBL operation is unlikely to achieve or sustain benefits.

## **E. Management of Policy-Based Lending Evaluations**

### **1. Coverage**

41. Since policy reform provides an enabling environment for other operations and for the economy as a whole, disproportionate coverage in evaluation should be provided. PBL evaluations provide important lessons for evaluations of country assistance strategies and programs. They should therefore be selected and timed to contribute to country program evaluations.

42. Individual PBLs should be evaluated in those cases where the outputs and outcomes can logically be separated from contemporaneous and subsequent PBLs, and when it is likely that the evaluation will generate lessons useful for replication or avoidance in similar operations. The main advantage of individual PBL evaluations is that this allows the evaluator to focus in-depth on the lessons of a given operation and to provide a well-reasoned performance rating.

43. Evaluations of a series or cluster of related PBLs can draw on cross-PBL experience to distill important lessons for formulating long-term programs of development policy assistance. PBLs should be evaluated in clusters or a series of PBLs in a sector or country over time when this provides a better measure of the policy reform assistance effort. Where a series of PBLs are evaluated together, it is recommended that they still be rated individually.

44. TA contributes to the capacity of a government to lead and manage the policy reform process. It is provided as part of an overall package of policy reform assistance support. Since both TA and PBLs contribute to policy reform performance, it is recommended that

- the performance of PBLs and associated TA should be evaluated and rated together.

### **2. Fostering Usage and Accountability**

45. Only if they are actually used do PBL evaluations have any value. To foster usage and accountability for delivering results, it is recommended that

- measures should be adopted to make evaluation products more user friendly.

Synthesizing and summarizing PBL evaluation findings in thematic, country, or sector reviews; making PBL evaluation reports shorter and more readable; disclosing them to the public; involving MDB boards, government, and management in the selection of thematic PBL evaluation topics; preparing PBL evaluation summaries for public dissemination; presenting PBL evaluation products verbally to boards and other key users; summarizing findings and lessons learned from PBL evaluations on in-house data bases; preparing press releases and briefings on PBL evaluation products; summarizing PBL evaluation findings in annual evaluation reports; hosting conferences and training sessions on PBL evaluation findings and methodologies; and encouraging public consultation on PBL evaluation methods and findings are recommended approaches to stimulate utilization of evaluation findings. Improving the linkages between PBL evaluation and performance accountability can also help to enhance the utilization of PBL evaluations, including lessons.

### **3. Disclosure**

46. Engaging stakeholders in the evaluation process can build trust, broaden ownership, and deepen understanding of the reform process. It is recommended that

- early public disclosure of PBL evaluation objectives, staffing, and approach should be encouraged to foster stakeholder participation and input in PBL evaluation.

47. Several MDBs are disclosing the findings of self- and independent PBL evaluation. This has helped to build confidence in the evaluation process and has contributed to a broader understanding of the process of policy change. It is recommended that

- there should be public disclosure of the main PBL evaluation reports and findings.

Only in exceptional cases, when disclosure of evaluation results is likely to seriously compromise the process of policy change, is some measure of confidentiality warranted.

## **Appendix: Subcriteria for Evaluating Policy-Based Lending Operations**

What follows is a suggested list of subcriteria. This is neither a comprehensive nor a minimum checklist. The subcriteria listed here have been found to be important determinants of policy-based lending (PBL) performance in multilateral development bank (MDB) evaluations and policy research. They can be used to select and define the criteria employed in evaluating specific PBL operations (PBLs). An evaluative judgment is required to assess the degree to which each chosen subcriterion has been achieved.

### **A. Relevance**

1. Consistency with the country's overall development strategy
2. Consistency with the MDB assistance strategy
3. Importance of unwritten objectives (i.e., was there a pressing financing imperative, and if so, to achieve what?)
4. Importance of policy objectives addressed to meet critical development constraints (by category, such as macroeconomic management, structural reform, sector reform, private sector development, institutional development, human development, environmental reform, and infrastructure development)
5. Reform ownership
  - leadership commitment: (a) locus of initiative, (b) level of intellectual conviction among policy makers, (c) expression of political will by top leadership, and (d) efforts towards consensus building among various constituencies
  - stakeholder support: (a) program designed as win-win reforms; (b) stakeholder assessment underpinning reform; (c) conditionalities structured to have greater chance of success given stakeholder interests; (d) tensions between stakeholder interests and reform resolved during policy dialogue; (e) key interest groups and the public made aware of the purpose, objectives, and likely effects of reform
  - institutional capacity: (a) ability and willingness of key institutions to reform factored into design, (b) adequate incentives for key institutions to participate in reform, and (c) assistance provided to augment identified gaps in skills or knowledge
  - participation: quality of stakeholder analysis and contribution
6. Validity of the diagnosis
  - policy dialogue considered alternative policy options and their consequences
  - sufficient country knowledge base for diagnosis and recommendations
  - lessons from past reform experience assessed and incorporated
7. Program Formulation and design was relevant to achieving objectives
  - adequacy of the external financing case for the operation?
  - adequacy of the overall external financing arrangements?
  - extent and appropriateness of the medium-term framework?
  - adequacy of the program logic (identification of the goals to be achieved, the specific purpose of the operation, reform measures, their expected outputs, outcomes, and development results, together with the key assumptions and risks to performance all identified)
  - appropriate reform instruments selected (reforms properly sequenced to reach policy targets, internally consistent, relevant, precise, not overly complex or detailed, realistic/feasible, a manageable number selected, aimed at substantial reductions in major distortions, cover key elements, and have clearly defined targets and objectives)
  - adequate tranching arrangements (disbursement phasing matches financing requirements, reform phasing matches disbursement phasing)
  - impacts (economic, poverty, social, environmental, and institutional) reflected in design

- social consequences assessed, and suitable mitigation measures incorporated
- performance risks (both internal and external) adequately identified, and suitable strategies for managing risk incorporated
- reform is coordinated with and coherent with other assistance efforts
- suitable implementation arrangements (well-functioning economic management team in place, high-level representation, burden sharing across ministries, and adequate incentives for affected parts of government to buy into the reform effort)
- adequate technical assistance (TA) available (to fill knowledge gaps, contribute to capacity building, and support program implementation)
- time frame realistic for reforms to be completed, given institutional and other constraints
- best option among alternatives selected

8. PBL could be readily evaluated

- targets well defined, linkages traced, baseline values provided, and performance targets specified
- reporting, monitoring, and evaluation responsibilities assigned and funding provided
- knowledge gaps identified, and actions identified for securing information needed for decision making

**B. Efficacy (Achievement of Purpose)**

1. Extent to which main program objectives were achieved

- by main written objective (i.e., desired outputs and outcomes achieved)
- by main unwritten objective (i.e., desired outputs and outcomes achieved)

2. Extent to which the political economy context was conducive to the achievement of program objectives:

- high social cohesion/ low degree of ethnic fractionalization
- recently elected government/politicians will be in office long enough to see the benefits of reform
- degree to which a perceived crisis has precipitated reform
- lack of large special interests represented in parliament
- public awareness and support obtained for PBL-supported reforms

3. Extent to which PBL financing and contemporaneous macro-reforms contributed to desired macro-outcomes

4. Extent to which PBL contributed to fulfilling key policy goals (i.e., macro stabilization, improved public investment, expenditures and domestic resource mobilization, financial sector reform, debt sustainability, private sector enabling environment, governance and public sector management, competitive pricing and marketing, labor legislation, sector regulation)

5. Extent to which PBL contributed to fulfilling key financial goals and objectives (i.e., financial restructuring, financial viability of distressed banks and state enterprises, enhanced cost recovery)

6. Extent to which PBL, and associated TA, contributed to fulfilling institutional development objectives (i.e., contributed to capacity for economic management and efficient and effective regulation, capacity of the executing agency to implement reforms)

7. Extent to which social objectives were met (i.e., poverty reduction, protection of vulnerable groups, reducing income disparities, community development and participation, gender equality, access and quality of public services, nutrition and food security, health improvement, participation, and empowerment)

8. Extent to which environmental objectives were met (i.e., improved natural resource management; maintenance of biodiversity; maintenance of soil, air, and water quality; improved urban environmental quality)

9. Extent to which PBL reforms contributed to continuity of the reform process (i.e., did learning, policy research, and dialogue advance the reform agenda and help build the constituency for the next generation of reforms?)
10. Extent to which factors beyond the government's control influenced the outcome of the program (including world markets, natural calamities, cofinanciers, performance of consultants, war/civil disturbance)
11. Were other performance assessments reviewed and presented (including those whose findings contradict the evaluation)?

### **C. Efficiency**

1. Socioeconomic benefits of major reforms are, or are expected to be, substantial
2. Socioeconomic benefits exceed costs (i.e., debt plus main adjustment costs)
3. Costs were effective compared with social objectives achieved (i.e., debt assumed and adjustment costs were relatively low compared with achievement of social objectives)
4. Financing was provided in a timely manner
  - identification and appraisal expeditious
  - financing disbursed in line with external financing requirements
  - disbursements took place according to plan
5. Transaction costs of providing assistance were a cost-effective use of assistance
  - costs of identifying, appraising, and supervising the operation lower, by volume of assistance, than for contemporaneous investment operations
  - government burden (i.e., proportion of economic management team resources devoted to design and implementation) did not preclude attention to other policy matters
6. Net economic returns to the reform operation
7. Public expenditures made adequate provision to meet the government's portion of adjustment costs

### **D. Sustainability**

1. Absence of major policy reversals
2. Continued borrower commitment to objectives demonstrated through post-program implementation of related measures
3. Sociopolitical support for the PBL (including adversely affected groups)
4. Adequacy of institutional arrangements for implementing agreed upon reforms
5. Conducive political setting (i.e., stable and supportive)
6. Conducive macroeconomic setting
7. Degree of reform resilience (i.e., resilience to changing financial, social, economic, and political conditions)

### **E. Institutional Development**

1. Contribution towards improving national capacity (in economic management, civil service reform, legal and regulatory systems, sector development, and others)
2. Contribution towards improving agency capacity (in planning, policy analysis, skills upgrading, public awareness building and consultation, management, restructuring, decentralization, management of information systems, financial controls, financial restructuring, regulatory enforcement, and agency governance)
3. Contribution towards improving private sector capacity
4. Contribution to improving nongovernment organization and civil society capacity (including participatory attitudes and involvement of society in the reform process)

5. Contribution towards improving the governance of the policy reform process (i.e., transparency, checks and balances, public participation and accountability in the process of reform).
6. Extent to which capacity has been developed to manage and lead the reform process within core government agencies

## **F. Impact**

1. Anticipated and unanticipated (positive and negative) impacts identified
2. Impacts attributable to the PBL isolated from those caused by other factors
3. Counterfactual scenario identified and used in impact assessment
4. Before and after conditions identified and used in impact assessment
5. Key development impacts assessed, either qualitatively or quantitatively, including
  - economic impacts
  - impacts on poverty
  - impacts on sector activity
  - impacts on the environment
  - impacts on social conditions (gender, participation, and others)
  - impacts on political conditions
6. PBL's additional contribution to the policy reform process identified

## **G. Borrower Performance**

1. Quality of preparation
  - degree of ownership and involvement in identification and design
  - political support for reform secured
  - adequate institutional arrangements for program implementation
2. Quality of implementation
  - well-functioning economic policy management team
  - effective leadership and management of the reform process
  - supportive macropolicies and complementary structural and sector policies
  - adequate public outreach, disclosure, and awareness building throughout program
  - key staff assigned to implement reforms
  - executing agencies had adequate incentives to undertake reforms
  - satisfactory use of TA
  - suitable mid-course adjustments made where necessary
  - covenants complied with or waivers requested
3. Adequacy of monitoring, evaluation, and reporting

## **H. MDB Performance**

1. Quality of PBL at entry
  - appropriate degree of selectivity
  - grounding in recent economic and sector work
  - adequate economic and financial rationale
  - adequate risk assessment
  - incorporation of lessons identified
  - adequate institutional analysis
  - adequate poverty, social (including gender), environmental, and stakeholder analysis

- incorporation of monitoring and evaluation indicators and reporting procedures
2. Quality of MDB supervision
- extent to which supervision and implementation contributed to achieving desired reform:
    - degree to which supervision focused on achieving objectives (versus conditionality compliance)
    - degree to which civil society participation was fostered in implementation (i.e., awareness fostered, prior notice given before reforms)
    - problems identified during implementation were expeditiously assessed and resolved
  - flexibility demonstrated in suggesting and approving modifications to achieve desired policy outcomes and impacts
  - supervision used to foster institutional learning and awareness about policy reform
  - quality of TA and advisory services during implementation
  - adequate resources devoted by the MDB to supervision
  - extent to which MDB reporting was complete, candid, and accurate
  - attention paid to monitoring and evaluation data and processes
  - quality and timeliness of self-assessment (i.e., completion reporting)

Sources: World Bank Operations Evaluation Department, Project Information Form and Project Assessment Form forms (1993 to the 2001 versions); African Development Bank Revised Evaluation Guidelines, 2001; Inter-American Development Bank, 1999, Protocol for the Special Study on Policy-Based Lending Phase III (evaluation questions); Asian Development Bank (AsDB), 2000. Guidelines for the Preparation of Project Performance Audit Reports; Steven Tabor and Stephen Curry, 2004, Good Practices for the Evaluation of Policy-Based Lending by Multilateral Development Banks, prepared for the Evaluation Cooperation Group of the Multilateral Development Banks, Manila: AsDB.