



Public Expenditure Management

Review of Bank Materials for OED

David L. Bevan

Director-General, Operations Evaluation: Gregory K. Ingram
Acting Director: Ajay Chhibber
Manager: Victoria Elliott
Task Manager: Catherine Gwin, Lead Evaluation Officer, OEDCM

This paper is available upon request from OED.

2003
The World Bank
Washington, D.C.



**ENHANCING DEVELOPMENT EFFECTIVENESS THROUGH EXCELLENCE
AND INDEPENDENCE IN EVALUATION**

The Operations Evaluation Department (OED) is an independent unit within the World Bank; it reports directly to the Bank's Board of Executive Directors. OED assesses what works, and what does not; how a borrower plans to run and maintain a project; and the lasting contribution of the Bank to a country's overall development. The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank's work, and to provide accountability in the achievement of its objectives. It also improves Bank work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.

OED Working Papers are an informal series to disseminate the findings of work in progress to encourage the exchange of ideas about development effectiveness through evaluation.

The findings, interpretations, and conclusions expressed here are those of the author(s) and do not necessarily reflect the views of the Board of Executive Directors of the World Bank or the governments they represent.

The World Bank cannot guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply on the part of the World Bank any judgment of the legal status of any territory or the endorsement or acceptance of such boundaries.

Contact:
Operations Evaluation Department
Partnerships & Knowledge Programs (OEDPK)
e-mail: eline@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125
<http://www.worldbank.org/oed>

Public Expenditure Management

Review of Bank materials for Operations Evaluation Department

David L. Bevan

Department of Economics
Oxford University

david.bevan@economics.ox.ac.uk

October 16, 2002

Public Expenditure Management: Review of Bank materials for Operations Evaluation Department

1. Introduction

This paper reviews a sample of World Bank documents, as well as other materials for workshops, falling within the broad area of public expenditure management.

The materials were selected by World Bank staff, and are supposed to represent the best of the Bank's work. There are eighteen sets of material, organised into four groups:

- Research and analytic reports (six documents)
- Economic and sector work (five documents)
- Projects (four documents)
- Supporting materials for major events (three sets of documents)

The paper is organised as follows. Section 2 describes the evaluation criteria and scoring system provided by OED. Section 3 summarises the quantitative scores awarded using these, with the detail being tabulated in the annex. Section 4 briefly discusses some caveats about this quantitative assessment. Sections 5, 6, 7 and 8 provide brief qualitative assessments of the four types of material and section 9 concludes.

2. The quantitative assessment procedure

To enhance comparability with parallel studies of other Bank activities being conducted by OED, I was asked to use a common set of six criteria for evaluation and a common rating scale with four scores. These are:

Evaluation criteria

1. *Relevance*: Was the information, analysis and advice in the document relevant to the client's issue?
2. *Comprehensiveness*: Was the information, analysis and advice in the document appropriately comprehensive?
3. *Level of Knowledge*: Did the document provide the client with the best and most up-to-date knowledge on the issue (including from sources outside the Bank)?
4. *Clarity*: Were the ideas and recommendations in the document stated clearly?
5. *Objectivity*: Did the document present information in an objective manner (indicating, where pertinent, significant differences that exist in ideas and approaches)?
6. *Practicality*: Did the document provide practical advice to decision-makers?

Quantitative Scores

The responses to the six evaluation questions were scored on a scale of 1 to 4, with 4 being the highest.

1. Clearly inadequate/not at all helpful
2. Barely adequate/not very helpful
3. Adequate/fairly helpful
4. More than adequate/very helpful

3. Quantitative Assessment

The Annex to this paper contains a table for each of the eighteen sets of material ('document') reviewed. These tables contain the quantitative score on each question, a remark explaining the basis for that score, an overall qualitative assessment of the document, and a mean score.

The documents reviewed are:

Major research/analytical reports

1. A Contemporary Approach to Public Expenditure Management, Allen Schick, 1998
2. WDR for 1997 – “The State in a Changing World”.
3. Public Spending for Poverty Reduction, April 2001, the Poverty Sourcebook
4. Actions to Strengthen the Tracking of Poverty-Reducing Public Spending in Heavily Indebted Poor Countries (HIPC), IMF and World Bank, March 2002
5. Reforming Public Institutions and Strengthening Governance. Nov. 2000.
6. Governance Strategy Update for Board, March 2002.

Key Economic and Sector Work

7. Turkey PEIR, 2001.
8. Tanzania PER, 2000
9. Tanzania PER, 2002
10. Vietnam PER, 2000
11. Croatia PEIR 2001

Projects

12. Cambodia - Economic and Public Sector Capacity Building Project (IDA/R2002-0106/1) June 25, 2002
13. Uganda – PRSC 1, 2001 (P7442)
14. Albania – Public Administration Reform, 2000 (PAD – 20059)
15. Chad SAC IV, 2001

Supporting Materials for Major Events

16. Public Expenditure Analysis and Management Course, May 21-23, 2002
17. Public Expenditure Reviews and Poverty Reduction: Issues and Tools, Cape Town, Feb 17-20, 2002
18. PREM Learning Week

Mean Score of all Documents Rated		
Item	Item Score	Group Score
Major research/analytical reports		3.89
Contemporary Approach to PEM, 1998	4	
WDR, 1997	3.83	
Public Spending Chapter in Poverty Sourcebook, 2001	3.5	
Actions to Strengthen Tracking, 2002	4	
Reforming Public Institutions, 2000	4	
Governance Strategy Update, 2002	4	
Key Economic and Sector Work		3.83
Turkey PEIR 2001	4	
Tanzania PER 2000	3.67	
Tanzania PER 2002	4	
Vietnam PER 2000	3.83	
Croatia PEIR 2001	3.67	
Projects		3.67
Cambodia, Public Sector Capacity Building, 2002	4	
Uganda, PSRC, 2001	4	
Albania, Public Administration Reform, 2000	2.67	
Chad, SAC IV, 2001	4	
Supporting Materials for Major Events		3.00
PEAM Course, 2002	3.83	
Cape Town Workshop, 2002	3.33	
PREM learning week	1.83	

The following table summarizes the mean score for the various types of documents on the six questions and one type of overall assessment score, namely the mean of the scores on the six criteria described above.

Summary of Average Scores for the four types of document

Document Type	Relevant	Compre- hensive	Knowledge	Clear	Objective	Practical	Overall
Research and Analytic	3.83	3.83	4	3.83	4	3.83	3.89
Economic and Sector	4	4	3.60	3.80	3.80	3.80	3.83
Projects	4	3.50	3.75	3.50	3.75	3.50	3.67
Supporting Materials	3	3.33	3.33	2	3.33	3.33	3

Across the entire set of documents reviewed, the mean score was between 3.6 and 3.7 for all criteria except that of clarity, which averaged rather lower at 3.3. Of the one hundred and eight individual scores, eighty-five were scored at 4, thirteen at 3, nine at 2, and only one at 1. Put differently, only three of the eighteen documents scored 2 or below on any criterion.

It is clear from these scores that this reviewer believes that these documents are highly successful on these criteria. This partly reflects the fact that this seems an impressive body of work. It also reflects some inadequacy in the assessment criteria. This second feature is discussed first.

4. Some Caveats about the Assessment Criteria

Impact

It is a central and welcome tenet of the Bank's current philosophy that improvements in public expenditure administration and performance require in-country commitment and institutional reforms. In keeping with the shift in best-practice budget management towards a concern with outcomes, the appropriate way to review the Bank's efforts in this area would require some assessment of impact. It seems plausible that a document scoring very badly on these six criteria would be unhelpful in improving outcomes. However, it would be perfectly possible for a document to score highly across the criteria and still have little useful impact, while a document with a less illustrious score might be very effective. For example, it is commonly felt by Bank staff that increased in-country participation, and hence ownership, often leads to a loss of quality in the programme design and accompanying documentation. A constructive shift to more participation might then lead to a falling score but a more productive process.

The case for assessing the outcomes of the Bank's work in this area, rather than the quality of the documentation, has been argued eloquently within the Bank. See for example "QAG Evaluations of Public Expenditure Reviews FY99-FY00", 11-8-00, by Cheryl Gray, Geoffrey Shepherd and Vinaya Swaroop. There are clearly

difficulties in moving in this direction, especially in the context of desk-based studies, but the issue is important.

Quality

Even in assessing documentary quality, the current procedure is rather crude, with only four allowable scores. The distinction between fairly and very helpful, for example, might be sharp enough for a document with a narrow technical focus. However, most of the documents reviewed here are mammoth combinations of analysis, advice, recommendations, and actions. Typically, they contain a whole raft of component pieces, some of which are likely to be very helpful, some only fairly helpful, while others are probably counter-productive. Deciding how to score the composite entity is not just a matter of ‘head-counting’ these categories and calculating a mean, but involves a judgement of which parts of the programme ‘really matter’. This adds a second layer of subjectivity to the exercise.

There is another aspect of this point. As the Bank is very ready to acknowledge, capacity is weak in many client countries, and the Bank’s own activities are very demanding of this capacity. There is always the danger of the best being the enemy of the better. Hence a document might score highly for being clear, knowledgeable, objective, and comprehensive while really being very over-ambitious in the country context. This might be picked up under the criteria of relevance and practicality, but not necessarily. In its PEM analytic work, the Bank has been rightly keen to stress that the target should be ‘best-fit’, not necessarily best practice. However, judging whether a proposal is best-fit is more demanding than judging whether it is best-practice: it requires familiarity with the country as well as with the relevant technical literature.

Conclusion

These difficulties with the assessment criteria are rather wide-ranging, and are not amenable to any quick fix. Hence no attempt is made here to supplement them in the present quantitative analysis. In any event, the common criteria do have some informational content even in the present context, provided they are interpreted with care.

5. Major research/analytical reports

The Bank has been at the forefront of recent developments in the area of public expenditure management, and has made major contributions to several aspects of these. Examples include the shift to more outcome related budget processes, the recognition of the centrality of institutional, ownership, and governance issues, and the need to modify a concern for best-practice with a recognition of the constraints of feasibility and hence desirability of best-fit procedures. In all this, the Bank has made good use of developments in industrial countries, while viewing these through the lens of its development mandate.

This set of six documents provides a good sample of this output, and as the scores indicate, it is a very impressive body of work. This general shift, and especially the new emphasis on institutions and governance, is sometimes dubbed the ‘Second Generation’ approach to contrast it with the previous more narrowly technocratic approach followed by the Bank. The WDR 1997 provided a major and up to date

analysis of the broad issues involved in defining the role of the state, and marked a welcome recovery of the Bank from a neglect of public sector issues over much of the previous decade. The 1998 Schick paper provided a well thought through discussion as to how the philosophy of the WDR might be implemented in the budgetary context. The 2000 Strategy Paper, *Reforming Public Institutions and Strengthening Governance*, carries this issue of implementation design further forward, and in a wider context. The 2002 Strategy update reviews progress and lays out a detailed plan of campaign for further action. The important matter of the relation between public expenditure and poverty reduction and how to improve the linkage is analysed in the remaining two documents.

In addition to the constructive developments documented here, the Bank has also been quite impressive in maintaining a lively internal debate on how best to trade these various factors against each other, and how best to assess the success of various reforms. For example, papers not directly reviewed here (though some are included in the Cape Town workshop materials) include documents from OED, such as “Evaluating Public Sector Reform” (2001) and “Governance - the Critical Factor” (2001). Similarly there are documents from PREM itself, such as “Public Expenditure Management and Accountability – Evolution and Current Status of World Bank Work” (2001) and “Guidelines for the World Bank’s Work on Public Expenditure Analysis and Support (including PERs)” (2001).

The two main reservations about the developments advocated in the documents reviewed here both relate to complexity. (These reservations are shared to some extent in the types of Bank document noted in the previous paragraph.) First, while this evolution has involved some substitution (as for example in budget process) much of it has involved addition. The demands on Bank staff and more especially on their client government counterparts have been escalating sharply over recent years. (This was a point made by staff at the Cape Town conference.) There is a real problem here. For example, in respect of one of the areas of increased loading, institutional reforms and governance, in many client states we cannot rely on the old simplifying adage “if it ain’t bust, don’t fix it”. But “fixing it” is likely to be a very slow, demanding and difficult process.

The second reservation concerns the difficulty of achieving clarity in the face of so much complexity. The documents reviewed here are quite successful in this regard. However, there remains plenty of scope for improvement. One source of occasional weakness is a lack of clarity about who the intended audience is. As the movement to wider participation grows, and involves not just an extended group of government officials, but civil society and other stakeholders, this multiplicity of audience types is going to stretch the Bank’s powers of communication. There are signs of a failure to resolve this issue, inter alia, in the *Poverty Sourcebook*. It may be necessary to accept that there are several audiences and to focus on them separately. I haven’t seen anything in the Bank’s documentation, for example, which is as user friendly to the relative newcomer as DFID’s guidelines for its advisers “Understanding and reforming public expenditure management” (2001).

6. Key Economic and Sector Work

The five documents in this group were Public Expenditure Reviews or the newer variant, Public Expenditure and Institutional Reviews. These have had very mixed reviews from OED and QAG in the past, though they have performed much more strongly in the last few years. The five examples here all achieved high scores in the present evaluation and seemed to be good examples of their type. (It should be noted that for the most part this reviewer has not seen peer review or other internal Bank assessments of these documents.) In general, they made impressive attempts to tailor modern good practice approaches to the particular country contexts. In short, they represent good attempts to implement the types of reform explored in the analytic material. It is also well understood in the Bank that not only the outcome of the PER but the process itself must vary with the country context. The PREM network's "Guidelines" noted earlier distinguishes between three 'prototype' PERs, wholly in-house, Bank-led participatory, and joint- or client-led. In this typology, my inference would be that the Croatia PEIR was in-house, the Turkey PEIR and Vietnam PER were Bank-led participatory, and the two Tanzania PERs were joint- or client-led. Whether this attribution is correct or not, the typology and the underlying philosophy it embodies seem helpful.

There is one question about it that requires a little thought, however. 'Participation' is used in this general area in two rather different ways. The more limited sense, applicable to Vietnam and Turkey for example, is that the consultative process extends beyond the central economic authorities to government more generally (and possibly other donors). The more expansive sense, applicable to Tanzania, is that consultation also extends to the private sector, trade unions, parliament, NGOs, the media, and civil society in general. The issue to be considered is whether it is desirable and feasible to press for this wider participation in the case of a Bank-led participatory PER, or whether that has to wait on fuller client involvement. In the case of Turkey in 2001, pressing for wider participation might have been premature; the priority was to build a wider constituency for reform within the government. This seems less clear in the case of Vietnam. The major sensitive issue – whether the SOEs should be reformed within the public sector, or privatized – had been settled in the former sense, as the Government wished. The Bank (and Fund) had effectively decided to keep their powder dry in this matter. What remains in the PER includes measures on which it would be useful to have a wider set of views. Hence, while the Vietnam PER is often quoted as highly participatory, this does not seem altogether right.

7. Projects

The four projects were all devoted to various forms of capacity building in the public sector, though it might be stretching a point to describe the Uganda PRSC as a 'project' at all. Three of them seem well designed and tightly focused, with a very clear structure of goals, actions, and means of monitoring performance. Once again, they demonstrate a clear familiarity with the foundations laid in the analytic work, as well as a realistic approach to implementation. The fourth project is the Public Administration Reform project for Albania. This seems much less satisfactory. It again demonstrates a good familiarity with the analytic literature but is couched in generalities. While the far-reaching nature of the PEM problem in Albania is spelt out

clearly, it is much less clear precisely how the project will address this, short of throwing a handful of expatriate advisers at it.

8. Supporting Materials for Major Events

These ‘events’ constituted two workshops and a ‘learning week’. The scores are uneven. The Public Expenditure Analysis and Management Workshop (2002) seems to have been very well designed and structured and the presentations are mostly self-sufficient, up-to-date and very clear. While the same is true of the individual presentations prepared for the Cape Town Workshop on PERs and Poverty Reduction, the design itself seems rather messy, with the sequence of presentations weaving back and forth between thematic issues and case studies in a confusing way. There is also a stupendous amount of material provided for the latter workshop (75 documents), using a selection principle that is opaque, and providing no guidance or prioritization. Meanwhile, the presentations for the PREM week seem very uneven, and it is unclear what audience would benefit from them. The week seems to be an ensemble of parts and does not appear to have been designed as a whole.

The first of these workshops is already accessible on the Bank’s website and it is presumably intended that the others, and/or some future events will be lodged there too. In that case, it will be important to ensure that all the presentations – as opposed to most – are fully self-contained, and do not assume the physical presence of all members of the audience.

9. Conclusion

Subject to the various reservations noted earlier, the overall conclusion of this review is that the Bank’s Public Expenditure Management activities are in pretty good shape, and that the Bank’s aim of becoming a ‘Knowledge Bank’ is in good order in this area of its concerns. Four qualifications to this rather upbeat conclusion are:

- The scale, complexity and ambition of what is being attempted, particularly in the light of country capabilities, and how best to keep these in line with each other
- The need to study carefully whether this ‘second generation’ approach really improves the outcomes that are achieved, including comparative studies of where it has worked and where it has not
- More thought as to the very different ‘audiences’ the Bank is addressing, and what this implies for the design of documents, the format of presentations transferred to the web, and the way in which ‘paths’ can be set up and signalled through related materials
- To guard against the (occasional) practice of invoking the ‘second generation’ mantra without designing tight programmes that would implement the associated strategy