

ASSESSING THE QUALITY OF FY00 ESW

APPROACH PAPER

1. As part of its continuing efforts to help improve the quality of the Bank's operational work, the Quality Assurance Group will carry out a review of the Economic and Sector Work (ESW) tasks completed in FY00. No major changes are proposed in the methodology from last year's review because the methodology has been widely endorsed in the Bank and has worked well, producing credible and useful assessments.¹ Based on the experience and feedback from the last exercise, and recommendations of a focus group of representatives from the Regions and the Networks, a draft approach paper was circulated to the Regions and Sector Boards. This paper takes into account comments and suggestions received on the draft. A few changes have been made in the guidance questionnaire. Questions have been added about objectives, scope and relevance of the task that are important to the Quality-at-entry of the task. Also a question has been added on the purpose of the task. As a pilot, two questions have been added on poverty issues. The section on Bank processes has been reconfigured for greater clarity. In order to strengthen the learning function of the review, it is also proposed to supplement the random sample with tasks nominated by the Regions as best practices (details of this proposal are in section 9), and include an observer from Regions/Sector Boards on in-depth panels.

WHY QUALITY ASSESSMENTS OF ESW?

2. ESW is important to the Bank for a number of reasons: it underpins lending operations, provides the basis for policy advice to our clients, and advances the base of knowledge of the entire development community inside and outside the Bank. At a direct annual cost of about US\$ 83 million², ESW is a major component of the Bank's operational services to its client members. As the Bank moves towards a more knowledge-based assistance to client countries, it becomes more important that ESW be carried out as effectively and efficiently as possible.

3. The program of ESW Quality Assessments is designed to:

- *Promote accountability* for quality of ESW by providing (i) senior management with indicators of progress on improving ESW quality; (ii) staff and managers with credible, independent assessments of the quality of a sample of tasks.
- *Identify systemic strengths and weaknesses* in ESW and promote appropriate changes in policies and procedures relating to ESW, including its management and review; and
- *Learn from experience* and help identify examples of best practices to emulate .

The methodology (desk-based with limited client consultation) is not designed to provide a definitive judgment on any individual task. Therefore, by itself, results from any single assessment are not adequate for staff/manager accountability. Feedback during assessment could, however, help improve the final quality and actual impact of the assessed product.³

¹ QAG plans to do a comprehensive retrospective of the three QAG Assessments of ESW after the end of this exercise.

² FY00 Plan

³ At the concept stage, QAG support is available through the QER program for quality enhancement.

HOW WILL THE ASSESSMENTS BE DONE?

4. **Coverage:** The Assessments will consider only ESW. Within ESW, the assessments will leave out products categorized as conferences/workshops, consultations, or study tours (accounting for less than 2% of ESW budget). They will not cover other Analytical and Advisory Activities (AAA) such as aid coordination, technical assistance, PFPs / PRSPs or fiduciary ESW such as CPARs and CFAAs. These may be reviewed subsequently in separate assessment programs following piloting during the next few months.

5. **Criteria:** The assessments will focus on whether we are doing the right ESW, doing that in the right way, and whether we are using the ESW products and processes to achieve maximum impact. They will be guided by four broad criteria:

- *Objectives, Scope and Strategic relevance:* This section asks if it was the right ESW to do. Are the objectives well-defined? Are issues addressed by the task appropriate in view of the objectives? Is the task relevant to the country circumstances? Is it worth doing, given competing demand for operational resources? Is it timely? Are audiences well-defined?
- *Internal Quality:* This section looks at the workmanship quality of the task. Does the report clearly specify the issues being addressed? How good are the analysis and the supporting empirical evidence? Does the analysis lead to substantive conclusions? Are the recommendations operational and consistent with the conclusions? Are they realistic in the light of social, political and administrative constraints?
- *Clarity of Presentation:* Are the main lines of the argument, and the conclusions and recommendations clearly spelled out? Does it have an adequate summary? Is the report comprehensible and transparent to clients, peers and other general readers?
- *Likely Impact:* Is the report likely to have impact on policy, institutional reform and/or investment programs? Did the work involve participation by the client and/or others? What is the actual or expected impact on the client, on the Country Assistance Strategy, or Bank thinking generally and on the wider development community, including key stakeholders?

6. In addition, assessments will consider whether the costs and elapsed time are commensurate with the quality of the report and with its (expected) impact⁴; was the work well managed and were appropriate actions taken to change its objective or scope when necessary? The assessments will also examine the quality of the Regions' internal review processes aimed at ensuring the quality of individual items of ESW, for example, whether there was an adequate peer review and an effective management review process?

7. **Ratings:** The relevance and relative importance of the above criteria will, of course, vary with the type of ESW analyzed and the country context. The panelists will use their judgement in deciding on the relative weights to be given to the individual criterion. Based on the experience of previous two reviews of ESW, panelists have generally given high weights to internal quality and relevance, and lowest weight to presentation. A Guidance Questionnaire to facilitate such

⁴ The same guidance questionnaire will be used for all tasks irrespective of the cost of the task. In assessment of 'small' tasks, special attention will be paid to cost-effectiveness in light of the available resources, but also including sunk costs and unrecorded subsidies.

assessments in a systematic fashion is shown as attachment. Quality will be rated on a four-point scale as below:

<i>Highly Satisfactory (1)</i>	Best practice in several areas and no major deficiencies
<i>Satisfactory (2)</i>	A good solid piece of ESW without major deficiencies
<i>Marginal (3)</i>	Major deficiencies in some important areas
<i>Unsatisfactory(4)</i>	A broad pattern of deficiencies

The ratings will take into account that the depth of analysis that can reasonably be expected will differ between countries, and between different types (sectors and topics) of ESW. The unit of assessment will be an ESW with its own task ID. It is recognized that all tasks may not be stand-alone pieces. They may be sub-tasks of a larger task or else one of a series of small tasks delivered over a period of time. In such cases, the assessors will try to understand as much about the context of the task as possible and factor it into their assessments. Task teams will be encouraged to provide summaries of related task reports and other relevant documents to help in this understanding.

8. **Sample:** The assessment will be based on a stratified random sample of ESW tasks "completed" (recorded as *delivered to the client* in SAP) in FY00. The sample will be stratified by cost and Region in order to sample more from tasks that cost more, and ensure representation of all regions in the sample. Like last year, four strata by cost are proposed, (i) >\$500,000 - all tasks will be assessed, (ii) \$250,000–500,000 - 50% will be sampled, (iii) \$50,000-250,000 - 20% will be sampled, and (iv) <\$50,000 - 10% will be sampled. If necessary, the sample will be expanded to include a minimum of five tasks from each of the six regions. The Regions have promised to deliver 335 tasks during FY00. Based on last year's distribution of tasks by cost, it is estimated that the sample would include about 60 tasks and cover approximately 30% of the total cost for all the tasks. Again, based on the information from last year's sample, the sample size should permit a robust assessment of how well ESW dollars are being spent bank-wide (with less than 5% sampling margin of error at 95% confidence level) and fairly precise performance assessments for most Regions and Networks. It will also permit estimation of ESW performance bank-wide in terms of number of tasks with a sampling margin of error around 5% at 90% confidence level. QAG will share the sampling universe with all the Regions and Sector Boards prior to drawing the sample and also the sample as soon as it is drawn. Tasks managed by Team Leaders who have been subject to two or more QAG evaluations in the previous 12 months will be replaced by other randomly selected tasks.

9. In order to strengthen the learning objective of this review and provide more robust benchmark for the *Highly Satisfactory* ratings, it is proposed to invite nominations of ESW (delivered to the client in FY00) exhibiting best practice from the Regions. The Regions will nominate these tasks after consultation with the Networks. The nominated tasks will be assessed along with tasks in the random sample, though **they will not be counted towards the quality score**. The qualitative lessons from these nominated tasks will be used to enrich the synthesis report and those rated *Highly Satisfactory* will be recognized just as the ones in the sample. This is also expected to establish a system of rewarding excellent ESW outside the QAG sample. **At this stage it is proposed as a pilot** and could be expanded in future QAG assessments. To keep the costs of the pilot low, it is proposed that all Regions nominate a maximum of 2 tasks each. If

a Region would wish to nominate more than two tasks, they would copay half the cost of assessment of additional tasks. The nominations will be from the ESW delivered to the client in FY00 and should be sent to QAG within two weeks after the final sample for the present review has been drawn. The total supplementary sample is expected to be around 15 tasks.

10. **Timing:** Assessments will be made after the task has been recorded as *delivered to the client* in SAP i.e., after the final substantive internal review, formerly known as Green Cover. They cannot interfere, therefore, with standard (Regional and Network) review processes and responsibilities. Indeed the value added by these reviews is itself an element of the assessment. It is hoped instead, that the patterns of weaknesses (and strengths) identified by the assessment will help to improve further the quality of future ESW.

11. The timing of the assessments, usually within 6 months after completion of an ESW task, also means that in many cases, it will not be possible to judge the actual impact of ESW. Preliminary judgements on likely impact may be possible based on the quality of the report, participation by or consultation with the client or stakeholder during its preparation, and actual or planned dissemination and follow-up. A more thorough assessment of impact would clearly remain the focus of self-evaluations by the Regions (through ACSs or otherwise) and Networks, and OED's Country Assistance Reviews and Sector Impact Evaluations.

12. **Rapid Assessments:** The assessment program will continue to be a two-stage process. In Stage I, two "rapid assessors" will assess each sampled task, using the guidance questionnaire as the framework.⁵ They will consult on their preliminary ratings and findings, including strength and weaknesses of the ESW task, and prepare a list of issues for discussion with the Task Team. The two assessors will finalize their consensus ratings and report after meeting only the Task Team normally, but Managers could join the discussion if they so wish. If the rapid assessors agree on a *satisfactory* (2) overall rating, they will inform (through QAG) the Task Team Leader/line managers of their assessment (report and the questionnaire), and will be available to discuss it further at their request. Where assessors intend to give an overall *satisfactory* rating to the task but a lower rating to the managerial performance, they will do so after consulting the concerned managers. If the initial overall rating is *less than satisfactory* (3 or 4), the task will normally go to Stage II for an in-depth assessment by a customized panel, unless QAG and the Regions concur that a panel would not add value. In addition, in all cases the rapid assessors may recommend to QAG to convene a panel in case of ESW tasks with potential for more learning. Rapid assessments (costing an average of 3 senior staff days per task) will thus serve as a screening device for the in-depth panel assessments (costing about 8 days per task). The proposed approach will ensure that no task rating of less than satisfactory is given without adequate consultation with the staff and managers directly responsible, or further panel review.

13. **Panel Reviews:** It is envisaged that about 25 ESW tasks will be subject to panel review (about 1/4 of the random sample and 2/3 of the nominated tasks). As with rapid assessments, the in-depth (panel) assessment will focus on ESW task quality, but with a different intensity especially with regard to consultation with the task team and management, and the client. The draft report of the panel will be discussed with the regional staff before being finalized. In addition, the regions (under the coordination of the Chief Economist) will be given the opportunity to attach a one-page rejoinder to the final report focusing on areas of divergent views (and reasons why) as well as on proposed follow-ups.

⁵ The team may be strengthened by adding more assessors if the nature of the task demands it (for example cross-sectoral tasks).

14. With a view to support staff development and disseminate QAG methodology in the Regions and Networks, in-depth panels will have an *observer* (a relatively younger staff) nominated by the Regions/Networks. QAG will collect feedback from the observers on what they learnt and also their impressions about the process.

15. **Inputs to assessment:** All assessments, both rapid and in-depth, will be primarily desk reviews. They will be based on a review of the task file, from the concept through the draft stage to the task completion report including comments by peer reviewers and others, minutes of review meetings specifying the decisions reached, and summaries of discussions with the client. The task teams will be encouraged to provide a short note describing the context of the task in terms of previous ESW, and other activities in the Bank and country. Where possible Stage I assessors will contact counterparts in the client countries in order to strengthen their judgment on *likely impact* (particularly when the task is process-oriented or the final product is to be produced by the counterparts). In-depth assessment will complement file review by extensive consultation with the task team, peer-reviewers and managers (Country Director and Sector Manager) concerned. Panels will also make an effort to obtain feedback from clients and other stakeholders to get an early reading on the impact of the task. To that end, they may obtain feedback from the offices of Executive Directors, the relevant IMF staff, or talk to client delegations that happen to be visiting Bank headquarters. However, as noted earlier, given the timing of the assessments and the absence of field visits, the focus of assessment will necessarily be on the quality of Bank performance and workmanship.

16. **Selection of Assessors:** The assessments will be carried out by experienced, credible, senior staff, from outside the Region concerned. Panel members and rapid assessors would normally be level GH or above, or equivalent consultants. QAG will select panelists or rapid assessors from lists of eligible staff and consultants to be established in conjunction with the Networks and the Regions. The eligibility lists would indicate country and sector/topic specialty, if any. To avoid any potential conflict of interest, anybody who has been involved in a particular ESW task, as peer reviewer, manager, or otherwise, will not be able to serve as rapid assessor or panelist on that task. The rapid assessor team for a particular ESW task will normally consist of a generalist, preferably with knowledge of the country/sector/topic, and a sector/topic specialist. To improve consistency across the evaluated tasks, where feasible, QAG will attempt to involve the same assessor in assessment of several similar tasks. Panels will normally comprise three members: a country specialist, a sector/topic specialist, and a generalist. QAG will provide a *moderator* for panel assessments to ensure consistency of approach by panels across sectors and regions.

WHAT WILL QAG DO WITH THE FINDINGS?

17. **Assessment Reports:** QAG will issues three types of reports: (i) brief rapid assessment reports for the roughly three-fourths of all cases where no panel is convened; (ii) more extensive panel reports; and (iii) an annual synthesis report. Rapid assessment reports with an overall “satisfactory” rating (2) will be distributed to the Task Manager, Country Director concerned as well as the relevant Sector Manager. All panel reports will, in addition, be sent to the Chief Economist, Directors Operations Support and the relevant Sector Board Heads. The synthesis report will provide an overview of the assessment findings, highlight patterns of strengths and weaknesses of ESW, and make recommendations for further improvements of ESW. The report will be sent to the Managing Directors, Regional VPs, Network VPs, Country Directors and

Sector Managers, as well as the staff at large. The final version will also be transmitted to CODE.

18. **Dissemination Channels:** QAG will also disseminate its findings through a variety of other channels, including:

- publishing best practice examples on the QAG web site.
- publicizing best practice examples through the Networks;
- drawing attention of Staffing Groups to skill-mix problems;
- organizing in partnership with WBI and Networks, targeted staff training workshops.

WHAT WILL IT COST?

19. Each rapid assessment of an ESW task is estimated to take an average of 3 staff days. Panel assessments will normally take 8 staff days. With the proposed program of some 75 rapid assessments and 25 panel reports (including those nominated by the Regions as best practice), the total assessment time would amount to 425 staff days. At an average cost of \$900⁶ per day for senior staff/consultants, this would come to around \$380,000. Including moderators' time and contingencies, the total costs for assessment of the FY00 ESW program will be about \$450,000 (this represents less than 0.5% of the annual ESW budget including management overheads).

20. QAG will cover direct costs of the assessments. These costs do not include the time spent by Regional staff in interacting with assessment teams. At about 0.5 staff day for each rapid assessment and 2 staff days for each panel, these costs will be around \$100,000. The staff are encouraged to monitor and report these costs through a budget code assigned by CRM.

WHAT WILL BE THE SCHEDULE OF KEY ACTIVITIES ?

Issuance of Approach Paper	April 13, 2000
Sampling of tasks completed during first 3 quarters of FY00	April 20, 2000
Sampling of tasks completed during last quarter of FY00	July 17, 2000
Nomination of best practice tasks by Regions	July 31, 2000
Rapid assessments and panels	May1 - Sep 29, 2000
Synthesis report	Nov 30, 2000

⁶ Assuming an equal number of staff and consultant days. A staff day is assumed to cost \$1200 and a consultant day \$600.