the standard, while avoiding a rating or "pass or fail" report.[31] While it is too early to judge the effect of such changes, this evaluation reinforces the view that the overall qualitative assessment and identification of key remaining gaps are the most critical elements and that the exercise should not be condensed solely to one of ratings, even if that is the aspect that market participants indicate that they value the most (see below).

v) **The governance structure for assessing standards is a little vague, but present arrangements for providing feedback work satisfactorily in practice**. The issue is—"who assesses the assessors?" In principle, this is the responsibility of the IMF and World Bank Boards for those standards assessed under the FSAP. In practice, members of the Board are not in a position (e.g., they are not provided with the necessary information) to make such judgments. The various standard-setting bodies, and their secretariats, do have the appropriate background but do not have governance responsibility for assessing whether the assessment exercises are proceeding satisfactorily; indeed, they do not even see those FSSAs that countries do not agree to publish. Nevertheless, discussions with the various standard-setting groups suggest that, in practice, there are sufficient informal and formal channels (including the Financial Sector Forum and IMF/World Bank staff participation in various technical committees) for adequate feedback to be provided on how assessments are being conducted. Our interviews with the various secretariats suggests a high degree of satisfaction with the results (see below).

IV. FSAP CONTENT

17. In judging the overall quality of the FSAP content, we relied upon two major sources of evidence. First, IEO assessors rated the content of the FSAPs for the 25 country cases on a 4-point scale according to various criteria: coverage and balance of assessments; clarity and candor of findings; as well as clarity, usability and prioritization of recommendations (see Table 2). These ratings on individual components were also used as inputs into an overall qualitative judgment on how well the FSAP assessment was integrated across the various sectors and with the overall macroeconomic picture. A high quality overall assessment is one that combines effectively the results from the various evaluation tools to present the main risks and vulnerabilities to the financial sector—i.e., those of macroeconomic/systemic significance—with an indication of criticality and consequence. Second, the surveys of various groups of stakeholders included questions on various aspects of the quality of FSAPs. We discuss first the overall quality of the FSAP assessment and then the articulation of findings and recommendations. Finally, some issues related to the "joint" IMF-World Bank nature of the FSAP are addressed.

[31] See *The Acting Chair's Summing Up: The Standards and Codes Initiative—Is it Effective? And How Can it be Improved?* July 25, 2005.

A. Overall Quality

18. **The overall quality of the FSAP assessments is high, although problems were encountered in a minority of cases**. The evaluation of the 25 countries reviewed shows, for example, good ratings for the overall financial sector coverage, the clarity and candor of findings, and the explanation of importance and consequence but with low ratings for a proportion of cases (typically about 10–20 percent) (Table 2). The survey of the authorities supports these conclusions, indicating a strong satisfaction with the adequacy of coverage and depth of analysis. Survey results for Article IV mission chiefs tend to agree with those views, but a larger share indicated some dissatisfaction with the results; in particular, about one-fifth of mission chiefs indicated dissatisfaction with the depth of analysis in FSAPs (see Figures 4a and 4b).

Figure 4a. Adequate Coverage of the Financial Sector

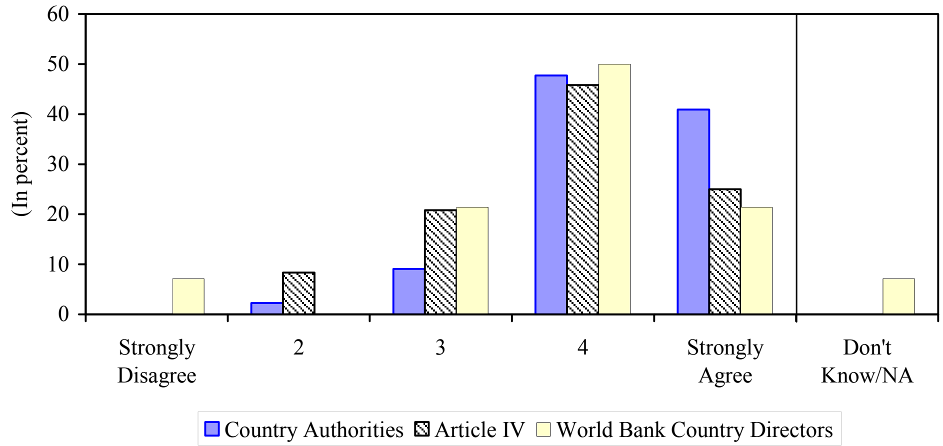
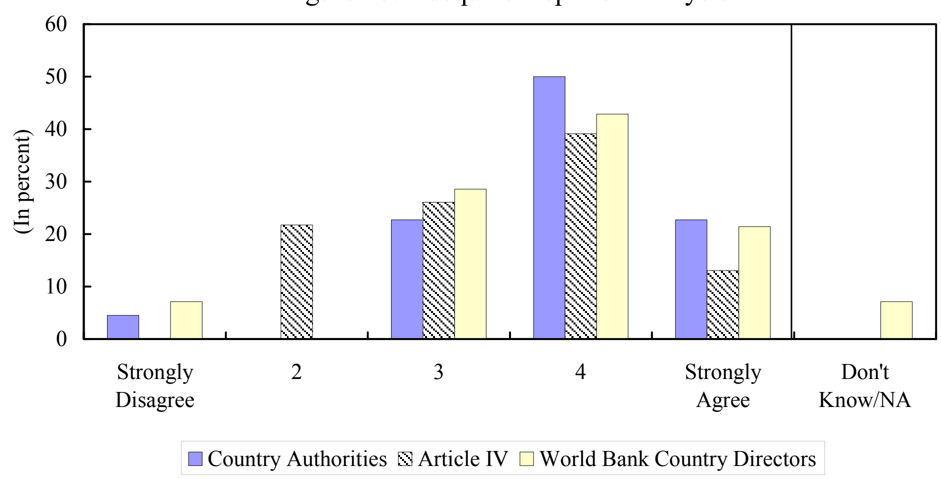


Figure 4b. Adequate Depth of Analysis



Source: Q5.1 and 5.2 of the survey of country authorities; Q7.1 and 7.2 of the survey of Article IV mission chiefs; and Q7.1 and 7.2 of the survey of World Bank Country Directors.

19. **The integrated approach to financial sector assessment does offer considerable advantages, which have been utilized in many but not all cases**. Such a "comprehensive" approach combines a variety of assessment instruments, coverage of the overall financial sector, and an analysis of the interaction between key macroeconomic risks and the financial sector in a manner that the sum is greater than the individual analytical components. A qualitative assessment of the 25 case studies suggest that in about 60 percent of the cases, this overall integration has been handled well (see Box 2 for examples). In about 20 percent of cases, the quality of the integrated assessment is broadly adequate but some gaps could have been filled by a better integration of the various components (e.g., in the case of Egypt's FSAP, important qualitative findings in the BCP assessment as well as data limitations impairing the analysis should have figured more prominently in the assessment). Finally, in about 20 percent of the cases, there are significant gaps in the overall assessment. For example, the Philippines FSAP does not make sufficiently clear the extent of weaknesses found in the banking sector nor their potential macro-systemic consequences, and there is no meaningful stress testing of the implications of key macroeconomic risks that were being actively discussed in surveillance reports.

20. **Analysis and integration of financial cross-border issues generally received limited attention**. FSAP stability assessments have generally been limited to the segments and risks of the financial system that have domestic implications, even when some external/macro risks were considered for the stress testing analysis (for example, the focus on the domestic consequences alone was especially notable in the Singapore FSAP). As noted above, FSAPs in countries with extensive financial sector cross-border activities have generally made limited inroad into the broader global and regional dimensions of those cases, with limited contribution to identifying and highlighting potential spillover channels and effects.[32]

21. **In a minority of cases, the overall assessment does not give a clear indication of the macroeconomic/systemic importance of vulnerabilities and potential consequences if key problems are not addressed.** The review of the 25 cases indicates that in about one-fifth of FSAPs there were significant shortfalls in the explanation of systemic importance and consequence of findings (Table 2). These mainly comprised cases where there was insufficient analysis of the criticality or urgency of vulnerabilities, the potential linkages and spillover effects to other segments of the financial system or corporate sector, and the macroeconomic impact and potential policy implications. Although it is not possible to

---

[32] Part of the work on financial cross-border issues is expected to take place in the context of regional FSAP exercises, especially for cases involving currency unions. A regional FSAP has recently been conducted for the Eastern Caribbean Currency Union countries, while similar exercises are planned or being discussed for some African and European countries. Given the very recent nature of the experience with regional FSAP exercises it is too early to assess their results.

assess statistical significance with our sample size, later vintages of FSAPs appear to have improved on the reporting of macro/systemic importance and consequence of findings.

Box 2. The Comprehensive Approach of the FSAP: Country Examples

A key potential value added of the FSAP is that it takes a comprehensive approach which is expected to result in an overall assessment that permits a greater understanding of the financial sector than would be possible through the separate assessment of specific components.

The FSAP is supposed to be comprehensive in several respects: the type of assessment instruments it applies (stress testing, FSI, ROSCs, etc.); the coverage of the financial system (banking, insurance, securities markets, payments systems, etc.); the analysis of interplays between macroeconomic and financial sector trends and policies; and in the identification of interactions with other economic sectors. This integrated approach is expected to strengthen the ability to recognize and analyze sectoral and macroeconomic linkages. Understanding these linkages in turn permits a fuller comprehension of risk and vulnerabilities, and the identification of potential policy options, complementarities and sequencing/prioritization needs.

While none of the 25 country cases examined in depth was "best practice" in all respects, there are many examples where the integrated approach has yielded results. Without trying to be exhaustive, the following examples can be mentioned:

- Cases in which macro/stability issues have a clear linkage with the financial sector. The case of Japan shows an FSAP where findings are embedded into a 4-pillar macro framework (broader and faster financial reforms, accelerated corporate restructuring, more aggressive monetary policy, and medium-term fiscal consolidation). Linkages and synergies of reforms are presented as part of the overall assessment, for example, on the need to address jointly corporate and banking reforms, and in highlighting the adverse effects of protracted low nominal interest rate on incentives to restructure bank portfolios.
- Cases in which the interlinkages among different markets are clearly analyzed. For example, the Chile FSAP identifies that pension funds' investment limits are creating scope for pension funds to provide a stable source of funding to the banking system. Similarly, the FSAP for Kazakhstan identified the connection between weak banking supervision and a structure of ownership linked to de facto conglomerates owned by some government officials that took control of recently privatized public enterprises.
- Cases in which the comprehensive analysis provides the elements for the design of a coherent program of structural reforms. For example, in the case of Mexico, the FSAP provides a comprehensive sequencing of necessary reforms in the capital market, including corporate government, institutional developments and banks' crisis resolution mechanisms.
- Cases in which incipient deepening of financial segments with inadequate regulatory and supervisory frameworks can have potential stability implications. The cases of Costa Rica and India highlight the challenges involved in countries where the financial system is evolving from one with a pervasive and commanding presence of the public sector to one where private sector participation takes a more prominent role. The challenges are partly associated with the need to sequence appropriately changes in organization and incentives in the financial sector with the necessary transformation of the regulatory framework .

22. **The effectiveness with which FSAPs addressed both stability and development issues in an integrated manner varied substantially and appears to have depended in part on the nature of the development issue.** While overall judgments by IEO assessors on the balance between stability and development issues in the 25-country sample were generally quite favorable (Table 2), FSAPs were more successful in handling some types of development issues than others. When the issue was one of reforming existing financial systems to promote growth, there tended to be a close association between the development

and stability aspects and FSAPs often handled these issues well. For example, recommendations on shifting from a public sector dominated banking system and a relatively closed capital market—in, say, India or Costa Rica—were primarily motivated by the goal of faster growth and development, but FSAPs rightly noted that managing the transition in a manner consistent with financial stability would require a carefully sequenced approach, including a strengthened supervisory framework. However, when it was a question of promoting the development of largely non-existent financial sectors, or encouraging the provision of financial services to underserved or excluded groups, the integration between the two aspects was generally handled less well. Indeed, whether the FSAP is the best vehicle to address such types of development challenges remains an issue (examples where the integration of such issues was not handled well include South Africa, discussed further below, and Kazakhstan).[33]

B. Articulation of Findings and Recommendations

23. **The main findings of the FSAP were generally presented in a reasonably candid manner** (in both the FSAP aide memoire and the more widely circulated FSSA), **although often couched in cautious language**. The detailed reviews of the 25 country cases and the broader survey results both indicate a relatively high rating on this category (Table 2). However, in the view of the IEO assessors, the language used was often very cautious and a franker presentation of key messages would have been useful. Officials of the various standard-setting organizations made a similar point, stressing that they found the overall messages of FSSAs highly informative but often couched in overly technical and oblique language (which some commentators referred to as "Fundese").

24. **In most cases, the main recommendations were clear and well linked to the findings**. (see Table 2 and Figures 5a and 5b).

25. **There were significant shortcomings in the prioritization of recommendations in many cases.** The ratings of IEO assessors for the 25 in-depth cases show problems with prioritization in over 40 percent of cases (Table 2). Similarly, only half of Article IV mission chiefs think that FSAP recommendations for their countries were well-prioritized (Figure 6); as will be discussed in the next section, this factor appears to have had a significant influence on the effectiveness of subsequent follow-up on financial sector issues in IMF surveillance. Difficulties with prioritization were more of an issue in countries where the FSAP assessment suggested the need for an extensive financial sector reform agenda (e.g., Ghana, Kazakhstan, Philippines), but it is precisely in those cases that effective prioritization is most important. Some, but not all, later vintages of FSAPs appear to have improved on the prioritization of recommendations. One recent "good practice" example of effective prioritization is the FSAP Update for Armenia (April 2005).

---

[33] The forthcoming OED report discusses further the effectiveness of FSAPs in assessing financial sector development issues.

Figure 5a. Recommendations were Clear

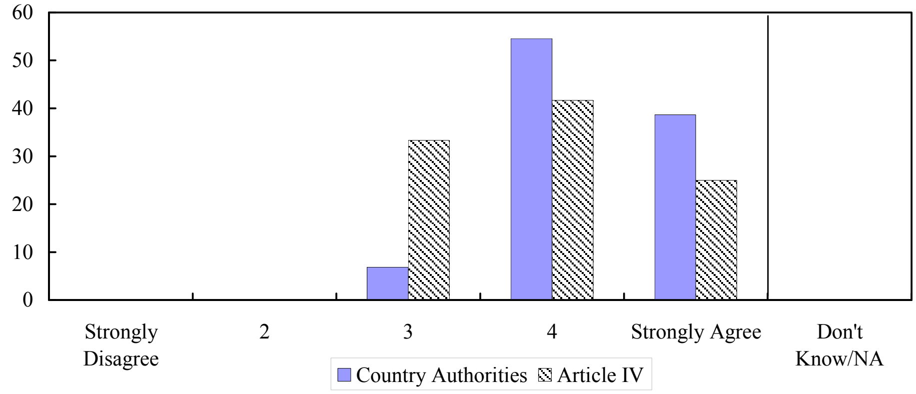
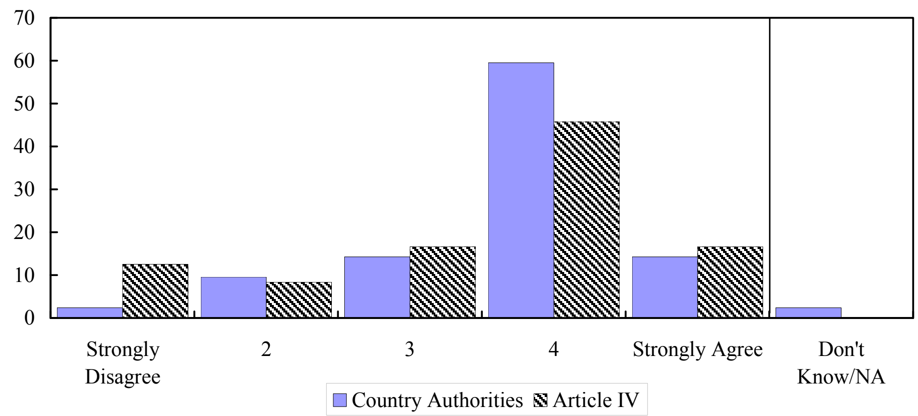
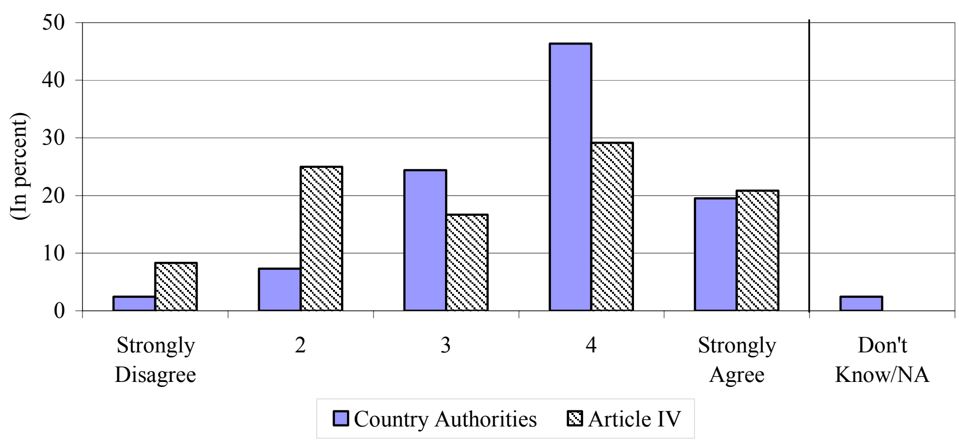


Figure 5b. Recommendations were Candid



Source: Q9.1 and 9.2 of the survey of country authorities; and Q11.1 and 11.2 of the survey of Article IV mission chiefs.

Figure 6. Views on Whether Recommendations were Prioritized



Source: Q9.3 of the survey of country authorities and Q11.3 of the survey of Article IV mission chiefs.

26. **While there was no major loss in candor across earlier stages of the FSAP process, candor was sometimes lost at the critically important stage of integration with Article IV surveillance reports**. The detailed review of documents produced at different stages of the FSAP process in the 25-country sample suggests that there was no significant loss of candor in the messages between the FSAP aide memoire and the FSSA. Moreover, there did not seem to be any significant difference in the level of candor between published and unpublished FSSAs. The informal presentations made to the authorities at the end of FSAP missions (PowerPoint presentations, etc.) that we were able to review were generally blunter, with more market-sensitive information, than any of the written assessments. This approach seems appropriate.

27. There was, however, a "loss in translation" in a number of cases between the messages of the FSSA and those incorporated in the staff reports for Article IV surveillance. This is critically important: as will be discussed further in the next section, when this happened the Board discussion tended to focus on the issues discussed in the Article IV report, crowding out problems that were flagged only in the FSSA, even though the latter report was also available to the Board. The following factors appear to have influenced how well key FSSA messages were integrated into the Article IV report:

- *Degree of country ownership.* In those cases where the authorities viewed the FSAP exercise as an opportunity to provide an independent judgment on weaknesses in financial sector policies and institutions and to catalyze reform plans, there tended to be a strong coincidence of interests in having the FSAP messages forcefully emphasized in the Article IV dialogue and staff reports (e.g., Costa Rica, Chile, and Mexico). However, when the authorities disagreed with key conclusions of the FSAP team, presentation of these conclusions in surveillance reports was often much more muted. The strongest example of this among our case studies was the Dominican Republic and reflected, in part, a failure of the internal review process to ensure surveillance reports reflected the key FSAP messages (see Box 3). Survey results support the conclusion on the importance of country ownership and suggest that a desire to avoid sending adverse signals was also important (see Figure 7).

- *Degree of integration between the work of FSAP and area department teams.* Weak country ownership of FSAP conclusions did not result in a loss of candor when there was close agreement between the diagnoses of the FSAP and area department teams. For example, the authorities in Japan and, to some extent, Germany did not agree with some important messages in their respective FSSAs, but surveillance reports reiterated these messages cogently.

28. **The review of FSAP Updates broadly confirms the findings on initial assessments** (see Box 4). Updates use the standard toolkit, and similar drawbacks in implementation are encountered. The articulation of findings and recommendations still presents weakness in prioritization, and integration into Article IV surveillance pertains more to reporting than expanding the overall macro assessment.

Box 3. Dominican Republic: The FSAP and the Subsequent Financial Crisis

The Dominican Republic provides an important example of the limitations of the FSAP process since a major financial crisis occurred shortly after the FSAP exercise was completed. This raises questions about how effectively the FSAP diagnosed the vulnerabilities that led to the crisis and about how the IMF used the results.

The FSAP was undertaken in 2001–02, and the FSSA was discussed by the IMF Board in June 2002. In early 2003, a run on one of the largest banks—Banco Intercontinental (Baninter)—occurred, triggered by the discovery of massive fraud. The central bank initially provided substantial liquidity support but eventually intervened the bank, removing existing shareholders and management. Similar problems related to accounting malpractices and mismanagement surfaced in two other banks. These events cumulated into a major financial sector crisis, with an eventual cost estimated at between 14 and 17 percent of GDP.

The evaluation, drawing on a detailed review of IMF internal documents and extensive interviews, reached the following conclusions:

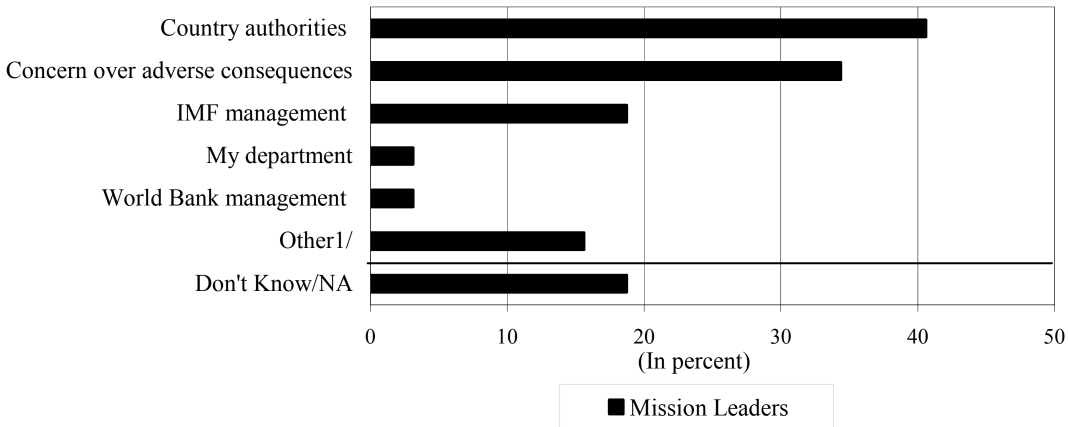
- The FSAP did not detect the immediate cause of the crisis, which involved the keeping of two sets of accounts by the banks involved. But FSAP exercises cannot be expected to detect accounting fraud, and are not a substitute for effective national audit and supervisory practices.
- The FSAP did diagnose severe and widespread vulnerabilities in the Dominican banking system, including an undercapitalized banking system, inadequate provisioning, overall weak compliance with BCP standards, and weak institutional capacity and judicial enforcement. Despite pressures from the then-government (who disagreed with the severity of the assessment), and the IMF area department (who recognized there were problems but thought the overall judgment too harsh in light of the Dominican Republic's then favorable economic performance), the FSSA presented to the IMF Board conveyed this assessment quite candidly (although some of the language was toned down from the aide memoire and, especially, from the initial PowerPoint presentation to the authorities by the FSAP team).
- The 2002 Article IV surveillance report failed to reflect the major warning signs flagged in the FSSA. It confined itself to an acknowledgement that the authorities were in agreement with the key findings of the FSAP (a statement which papered over many substantial disagreements). The staff appraisal referred to progress in reforming the financial system, without giving an indication of the huge challenges and dangers involved.
- The June 2002 Board discussion (of both the Article IV reports and the FSSA) largely followed the emphasis given in the Article IV staff report and did not focus much on financial sector issues. 2/ The FSAP deputy team leader was asked only a few technical questions at the Board meeting. 3/
- Thus, while the FSAP exercise was broadly successful in diagnosing many of the problems of the banking system (if not the extent of balance sheet problems hidden by accounting fraud), the surveillance process failed to utilize the assessment effectively. While the authorities began to implement some FSAP recommendations (such as adapting a new Monetary and Financial Law) with MFD assistance, the FSAP had little overall impact on the subsequent outbreak of the crisis. It is not possible to say whether a more effective integration of the FSAP with surveillance would have increased its impact, especially since the then-government had little ownership of the key messages and it was probably by then already too late to avoid the balance sheet problems at the heart of the crisis.

1/ A detailed discussion of the crisis is provided in the 2003 Article IV staff report.

2/ The subsequent Press Information Notice spoke of the importance of "further strengthening the banking system" and said the Board "… was encouraged by the progress made in reforming the legal and regulatory framework…" Since the FSSA was not published, the PIN represented the only public signal with regard to IMF surveillance of the financial sector and did not convey an adequate sense of the existing vulnerabilities.

3/ The FSAP team leader was from the World Bank and the deputy team leader from the IMF.

Figure 7. Reasons for Non-Candid FSAP Recommendations



Source: Q10 of the survey of FSAP mission leaders.
1/ Includes pressure from specific departments/individuals within the IMF or the World Bank.

Box 4. Assessment of FSAP Updates

Each of the 11 FSAP Updates completed in the post-pilot phase as of June 2005 was reviewed following a streamlined template based on the one used for the in-depth 25-country sample. 1/ The key messages from this review were as follows:

- While the scope of the various Updates has been implemented flexibly, in line with the Board's guidance—from a comprehensive reassessment involving a 16-person team in Colombia to a narrowly focused review of a few issues, involving a 2-person team, in Iceland—there continues to be only limited discussion of the rationale for the scope of Updates. (As noted earlier, the Ghana TOR is one exception.)
- Thus, it is difficult to see how each of the FSAP activities fits into an overall strategy for financial sector surveillance in each country.
- While most Updates conducted a new round of stability assessments, including stress tests, in most cases there was little improvement in methodological approach; thus, in a number of cases, data limitations still forced a highly simplistic approach. This raises questions as to whether a greater ex ante assessment of changes in data availability might have concluded that updating such tests was not a high priority use of resources.
- Updates were generally effective in conducting an in-depth tracking of implementations in the specific sectors covered in comparison to other mechanisms of surveillance.
- In general, the review suggests that reasonably comprehensive Updates—encompassing all sectors of significant macroeconomic importance—were needed to provide an overall assessment of progress in implementation to address identified vulnerabilities and remaining challenges.
- The limitations on what can realistically be expected from Updates of different scope and depth (e.g., that narrowly-focused Updates cannot be expected to provide an in depth assessment of progress in sectors that fall outside of its scope) are not adequately signaled. Stronger "health" warnings on these limitations and the necessary qualifications to any conclusions are still needed.
- The degree of integration of findings into surveillance reports appears to be broadly similar for Updates as for full FSAPs. The principal messages are reported but there is often little integration into the overall surveillance assessment. Slovenia is a "good practice" exception.
- Inadequate prioritization of recommendations remains a problem in most cases.

1/ The 11 Updates completed in the post-pilot phase are: Iceland, Ghana, Slovenia, Kazakhstan, El Salvador, Senegal, Colombia, Peru, Armenia, Hungary, Uganda. Under the category of Updates, there are different exercises, ranging from comprehensive to more focused. The term "update" is now used for this entire range of exercises.

C. The Contribution of the "Joint" (IMF-World Bank) Nature of the FSAP

29. The principal rationales for making the FSAP a joint IMF-World Bank initiative were that, in light of the overlapping mandates of the two institutions on financial sector issues and the scarcity of technical expertise, considerable potential synergies could be attained by addressing stability and development aspects in a comprehensive manner and that combining the respective expertise of the two institutions would produce a more integrated analysis and set of recommendations. It was expected that these gains would outweigh any additional costs of coordination. The evaluation assessed whether these potential synergies are being achieved in practice. The primary focus is on the implications of the joint initiative for the IMF, but we also draw upon the conclusions of the parallel OED evaluation.[34]

30. **Organizing joint teams that include both IMF and World Bank staff members (as well as outside experts) has contributed significantly to the depth of analytical expertise and credibility of the findings in many, but not all, cases**. Positive examples such as Chile, Costa Rica, India, Mexico, Russia, and Romania reflect cases in which the two staffs either contributed specific expertise not available in the other institution (e.g., Bank staff often contributed substantial expertise on non-bank financial institutions, including insurance, and on corporate governance) or where Bank staff contributed in-depth country knowledge gained in the context of other sectoral or loan negotiation activities. Indeed, these examples of positive synergies were not limited to cases where "development" issues received a major emphasis in the FSAP but were also present in a number of cases where the focus was almost exclusively on stability issues, including the strength of the supervisory system (e.g., Dominican Republic, Jordan, and Slovenia).

31. **Discussion of the relative weight to be given to stability and development issues was generally inadequate in earlier cases but there have been some improvements over time, both in the TOR initiating each FSAP exercise and in the FSAP aide memoire and FSSAs**. This trend reflects the increasing emphasis on streamlining and prioritization following the 2003 review. Clear understandings with the authorities on priorities are critical. For example, the South African authorities wanted their FSAP to focus on an assessment of financial sector stability and the strength of the supervisory framework. Although financial development—especially how to expand the provision of financial services to the half of the population with little or no access—was a key policy issue, the authorities did not regard the FSAP as the most appropriate instrument for addressing such matters. Partly as a result, the sections of the report dealing with these issues were piecemeal add-ons and judged to be of limited value added by both Bank-Fund staff and the authorities. In contrast, a good example of well explained—and appropriate—prioritization is Ghana. The initial FSAP (2001)

[34] The results discussed here draw on the case studies for developing countries only. However, many of the FSAPs for industrial countries also address longer-term structural aspects and their implications for the stability and efficiency of the financial system (e.g., Germany, Japan).

focused primarily on stability issues because there were pressing issues to be addressed. The subsequent FSAP Update (in 2003) focused primarily on a strategy for financial sector development. In both cases, these priorities were agreed with the authorities. However, such examples are still the exception rather than the rule in the FSAPs for developing countries.

32. In contrast to the more comprehensive use of various indicators and assessment tools for financial sector stability, most FSAPs still present a more limited analysis of financial development issues including access to financial services. Tools for the analysis of such issues remain less well developed.[35]

33. **While the degree of emphasis on stability and development issues varied substantially (and appropriately so) across countries, our overall judgment is that the degree of integration between the two was generally quite good but with significant shortcomings in a minority of cases** (see Table 2). What constitutes a "best practice" approach to such integration? The standard is perhaps clearest in those cases where substantial reform of the financial sector, and related policies, is needed to remove longer-term impediments to growth but where the process of reform (i.e., getting from a to b) itself could entail substantial risks of increased financial sector stability. In such cases, we would expect to see a clear analysis of such potential risks and of strategies for minimizing them (although not necessarily a very detailed blueprint). Among our sample countries, the FSAPs for Chile, Costa Rica and India provide examples of such "good practice" approaches. In contrast, the Tunisia FSAP does not quite meet this admittedly high standard, (although subsequent IMF technical assistance did contribute to such a strategy) and the Philippines FSAP provided little effective integration of the two aspects.

34. **The evaluation suggests no evidence that the joint approach has led to a "watering down" of messages in order to achieve consensus between the Bank and the Fund**. Indeed, the in-depth country reviews indicated two cases (Dominican Republic and, to a lesser extent, Russia) where World Bank staff helped to resist pressures that arose within the IMF's internal review process to tone down the FSAP messages on some aspects.

V. How Well Has the IMF Used the FSAP Output?

35. The evaluation examined how effectively the IMF used the FSAP output in each of its three primary activities—surveillance, technical assistance, and program design.

---

[35] The 2003 joint review of the FSAP by staff from the Bank and IMF presented an annex that discussed issues of financial development. But an operational framework for the assessment of development issues is less advanced than for stability aspects. The OED evaluation reaches a similar conclusion.