Evaluation Report

Fiscal Adjustment in IMF-Supported Programs

International Monetary Fund • 2003

© 2003 International Monetary Fund

Production: IMF Multimedia Services Division
Cover: Martina Vortmeyer
Figures: Theodore F. Peters, Jr.
Typesetting: Alicia Etchebarne-Bourdin

Cataloging-in-Publication Data

Fiscal adjustment in IMF-supported programs—[Washington, D.C.]: International Monetary Fund, 2003.

p. cm.

At head of title: Evaluation report.
Includes bibliographical references.
ISBN 1-58906-274-4

1. Fiscal policy. 2. International Monetary Fund. I. International Monetary Fund. Independent Evaluation Office.

HJ192.5.F47 2003

Price: US$25.00

Please send orders to:
International Monetary Fund, Publication Services
700 19th Street, N.W., Washington, D.C. 20431, U.S.A.
Tel.: (202) 623-7430 Telefax: (202) 623-7201
E-mail: publications@imf.org
Internet: http://www.imf.org

recycled paper

Contents

Foreword vii

Abbreviations ix

FISCAL ADJUSTMENT IN IMF-SUPPORTED PROGRAMS

Summary of Findings and Recommendations 3

Framework 3
Findings: Quantitative Aspects of Fiscal Adjustment 4
Social Spending and Social Protection in IMF-Supported Programs 7
Reforms in the Fiscal Area Under IMF-Supported Programs 9
Recommendations 11

1 Introduction 14

2 Fiscal Targeting in IMF-Supported Programs: Cross-Country Analysis 16

Relevant Considerations in Determining the Fiscal Stance 16
Factors Determining the Scale and Nature of Fiscal Adjustment 20

3 Fiscal Adjustment as Presented in Program Documents and the Internal Review Process 22

Fiscal Adjustment in Program Request Documents 22
The Internal Review Process Prior to Board Approval and During Program Implementation 25

4 Fiscal Performance Compared with Targets 30

Cross-Country Analysis 30
Flexibility of Fiscal Targets During Program Reviews 35
What Accounts for Large Fiscal Underperformance? Evidence from the Smaller Sample 39

5 Economic Recovery and Growth 42

Economic Recovery in the Program Period: Outcomes and Expectations 42
Optimism in Projecting Private Demand and Investment 45
Is the Fiscal Stance in IMF-Supported Programs Unnecessarily Contractionary? 47

6 Social Spending and Social Protection in IMF-Supported Programs 50

Has Donor Aid Been Underestimated? 50
Social Spending Under IMF-Supported Programs: Cross-Country Evidence 51

Role of the IMF in Connection with Social Expenditure and Social Protection 53
Conclusions 59

7 Fiscal Reforms in IMF-Supported Programs 61

Fiscal Reforms in Programs: An Overview 61
Progress in Implementing Reforms 63
Learning from the Past and the Role of Surveillance in Monitoring Progress 68
Conclusions 72

Boxes

3.1. How Well Do Documents Explain the Rationale for Fiscal Adjustment? 23
6.1. The 1997 Guidelines on Social Expenditure 55
6.2. How Public Hospitals in Ecuador Adjusted in a Time of Crisis 58
6.3. Protecting Critical Programs Is Not Costly When Programs Are Well Targeted 59
7.1. Public Finance Reform Areas 62
7.2. Ecuador: What a Determined Head of Tax Administration Can Do 64
7.3. The Romania Program: Diminishing Returns to Raising Tax Rates 68
7.4. Tax Reform in the 1996 Tanzania ESAF 69
7.5. Good Examples of Learning from the Past 72

Figures

2.1. Distribution of Programs According to the Magnitude of the Envisaged Change in the Overall Fiscal Balance (*T*–1 to *T*+1) 19
4.1. Distribution of Programs According to Differences in Fiscal Adjustment Between Original Targets and Reviews 36
5.1. Distribution of Programs According to Differences Between Actual and Envisaged Cumulative Growth over a Two-Year Period (*T* and *T*+1) 45
5.2. SBA and EFF Programs According to Differences Between Actual and Envisaged Investment Rates for *T*+1 46
6.1. Estimated Impact of a Two-Year IMF-Supported Program 54
7.1. Index Indicating Implementation Progress in 153 Fiscal Reform Measures in 15 IMF-Supported Programs 65
7.2. Progress in Implementing Fiscal Reforms in 15 IMF-Supported Programs 66
7.3. Index of Performance: Learning, Follow-Up, and Links Between Programs and Surveillance 70
7.4. Distribution of Cases According to Performance 71

Tables

2.1. Initial Conditions as Seen by Staff at the Start of the Program 17
2.2. Program Projections: Changes in Balances from (*T*–1) to (*T*+1) 17
2.3. The Direction of Change in Selected Macroeconomic Targets in IMF-Supported Programs as a Share of GDP 18
2.4. The Pace of the Envisaged Fiscal Adjustment 20
3.1. Degree to Which Program Documents Explain the Rationale, Magnitude, and Composition of the Envisaged Fiscal Adjustment 24
3.2. Average Public Debt Prior to and Following the Initial Program Year 26
3.3. Selected Topics Commented on During the Review of 15 IMF-Supported Programs 27

4.1. Changes in External and Fiscal Balances from (*T*–1) to (*T*+1) 31
4.2. Differences Between Actual and Projected Changes in Fiscal Balances 31
4.3. The Dynamics of Fiscal Adjustment 32
4.4. Percentage Distribution of Programs with Fiscal Shortfalls 33
4.5. The Composition of Fiscal Adjustment in Programs with Fiscal Underperformance 33
4.6. Changes in Government Balances in Large Episodes of Envisaged Adjustment from (*T*–1) to (*T*+1) 34
4.7. Distribution of Programs According to Revisions in Growth and Fiscal Balances 37
4.8. Revisions in Fiscal Balance Targets for *T*+1 37
4.9. Summary of Reasons for the Revised Fiscal Balance Target 38
4.10. Envisaged and Actual Fiscal Adjustment in Nine IMF-Supported Programs 39
4.11. Comparing Growth and Revenue Underperformance 40
4.12. Revenue-Related Structural Reform Measures in Selected Programs with Large Revenue Shortfalls 40
5.1. Experience with GDP Growth Prior to and During Program Periods: Annual GDP Growth 43
5.2. Programs Showing Deceleration or Negative Growth 44
5.3. Envisaged and Actual Two-Year Cumulative Growth Rates over *T* and *T*+1 44
5.4. Indicators of Growth Optimism for *T*+1 45
5.5. Investment Projections and Actuals Under IMF-Supported Programs, 1993–2001 46
5.6. Macroeconomic Balances in Stand-By and EFF Arrangements in Nontransition Countries 48
6.1. Public Sector Social Spending Indicators 52
6.2. Number of Countries With and Without Statistically Significant Results 53
6.3. Effectiveness in Identifying and Monitoring Social Spending in the Program Requests of 15 Selected Arrangements 56
6.4. The Ecuador Program: Imbalance Between Efficiency and Equity Measures Underpinned by Conditionality 57
7.1. Distribution of Areas of Fiscal Reforms and Those Supported by Conditionality 63

Appendixes

1 Table A1.1. Determinants of the Envisaged and Actual Fiscal Adjustment (*T*–1 to *T*+1) in IMF-Supported Programs 74
Table A1.2. Determinants of the Differences Between Envisaged and Actual Fiscal Adjustment 74

2 Code Book for Assessing the Need for Fiscal Adjustment 75

3 Table A3.1. Levels of Grants in a Sample of Sub-Saharan African Countries 77
Table A3.2. Changes in Levels of Grants 78
Table A3.3. Aid Flows Under IMF-Supported Programs, 1995–2001 79

4 Explanatory Variables and Methodological Issues in the Analysis of Social Spending in IMF-Supported Programs 80
Table A4.1. ARIMA Model with Control Variables and Endogenous IMF-Supported Programs 81
Table A4.2. Summary of Robustness Analysis 82

Table A4.3. Summary of Regression Results 83
Table A4.4. Control Variables for Social Spending 85
Table A4.5. List of Countries and Subsamples 86

5 Coverage of Social Issues in IMF-Supported Programs 88
6 Progress with Reforms in 15 Selected IMF-Supported Programs 96

7 Illustrative Selection of Technical Assistance Inputs to Fiscal Reforms in the Lead-Up to the IMF-Supported Programs in 15 Countries 97

8 Table A8.1. Effectiveness of Surveillance 99
Table A8.2. Article IV Reports Reviewed in the Preprogram Period 100

Bibliography **102**

Statement by Managing Director, IMF Staff Response, IEO Comments on Staff Response, and Summing Up of IMF Executive Board Discussion by Acting Chair

Statement by Managing Director **109**

Staff Response **110**

IEO Comments on Staff Response **113**

Summing Up of IMF Executive Board Discussion by Acting Chair **114**

The following symbols have been used throughout this report:

– between years or months (e.g. 2000–01 or January–June) to indicate the years or months covered, including the beginning and ending years or months;

/ between years (e.g. 2000/01) to indicate a fiscal (financial year).

"Billion" means a thousand million.

Minor discrepancies between constituent figures and totals are due to rounding.

Some of the documents cited and referenced in this report were not available to the public at the time of publication of this report. Under the current policy on public access to the IMF's archives, some of these documents will become available five years after their issuance. They may be referenced as EBS/YY/NN and SM/YY/NN, where EBS and SM indicate the series and YY indicates the year of issue. Certain other documents are to become available ten or twenty years after their issuance depending on the series.

Foreword

Fiscal adjustment is widely regarded as one of the core elements of macroeconomic design in IMF-supported programs, and has often been the source of controversy and criticism. The IMF has been criticized in various quarters for adopting a standard "one-size-fits-all" approach to determining the desired level of fiscal adjustment without taking into account country-specific constraints and circumstances. Concern has also been expressed that there is an excessive focus on fiscal austerity, which hurts the poor and may also produce a contractionary bias in IMF-supported programs.

This report examines these and other aspects of fiscal adjustment in IMF-supported programs based on analysis of a cross-section database covering 133 programs in 70 countries, supplemented by a more in-depth examination of program-related documents for 15 programs. The evaluation throws valuable light on many of the issues which have been the subject of controversy. The report finds that there is much more variation in the pattern of fiscal adjustment across programs than is generally assumed. For example, contrary to the general perception that IMF-supported programs invariably enforce austerity, it finds many instances where fiscal deficits were actually projected to widen and expenditures to increase as a percentage of GDP. The report also does not find evidence of a general contractionary bias leading to a slowdown in growth compared with precrises averages.

The report also finds evidence of weaknesses in program design in certain areas. There is a tendency to adopt fiscal targets based on overoptimistic assumptions about the pace of economic recovery leading inevitably to fiscal underperformance and frequent revisions of targets. The optimism about growth recovery in the short term is itself often the consequence of overoptimistic assumptions about the pace of revival of private investment when a more realistic assessment in certain circumstances could have justified the adoption of a more relaxed fiscal stance on contracyclical grounds.

The report also deals with the issue of the impact of fiscal adjustment in social sector expenditures, which are critical for the welfare of the poor. Cross-section analysis at the level of aggregate social sector expenditures does not find that these expenditures are lower than they would have been in the absence of an IMF-supported program. However, the in-depth country studies show that even when aggregate social sector expenditures are maintained, critical areas of expenditure most relevant from the point of view of the poor may be crowded out by certain components of expenditure such as wages and salaries. These adverse effects could be avoided at relatively small cost if ways could be found of protecting these critical expenditures during times of crisis.

The report makes specific recommendations on how to deal with these problems in the future, both in surveillance activity and in program design. The findings of the report and the recommendations were discussed in the Executive Board on August 29, 2003, and the reactions of the Board are summarized in the Acting Chair's summary, which is published along with the report.

The issues examined in the report are complex and often call for more data than were available in program documents. There is clearly a need for more detailed studies based on in-depth examination of individual cases. On its part, the IEO hopes to build on the results presented in this report in future evaluations where the same issues may arise. Hopefully, independent research by others will also contribute to a fuller assessment.

Montek S. Ahluwalia
Director
Independent Evaluation Office

Fiscal Adjustment in IMF-Supported Programs

This report was prepared by a team headed by Marcelo Selowsky and including Ali Mansoor, Steve Kayizzi Mugerwa, Alex Segura-Ubiergo, and Tsidi Tsikata, and was approved by Montek S. Ahluwalia, Director of the Independent Evaluation Office (IEO). Chapters of the report benefited from inputs from Patrick Conway and Ricardo Martin, and comments from Mario Blejer, David Goldsbrough, and Arnold Harberger. The evaluation also benefited from early comments received at a workshop organized by the German Foundation for International Development (DSE) (Berlin, July 2002), and one organized by the National Bureau of Economic Research (Cambridge, Massachusetts, October 2002). Experts in the field who contributed to the report include Gustavo Arteta (Ecuador), Rosario Gregorio-Manasan (Philippines), Cornel Tarhoaca (Romania), and Samuel Wangwe (Tanzania). Research assistance from Carolina Gutierrez de Taliercio, Mimi Tesser, Rouben Atoian, and Mwaffak Taib is gratefully acknowledged.

Abbreviations

CAS	Country Assistance Strategy
EFF	Extended Fund Facility
ESAF	Enhanced Structural Adjustment Facility
FAD	Fiscal Affairs Department
FMA	Fiscal management assessment
GDP	Gross domestic product
GEE	Generalized Evaluation Estimator
GST	Goods and services tax
HIPC	Heavily indebted poor country
IDB	Inter-American Development Bank
IEO	Independent Evaluation Office
LOI	Letter of Intent
MEFP	Memorandum of Economic and Financial Policies
MONA	Monitoring of Fund Arrangements database
NGO	Nongovernmental organization
OLS	Ordinary least squares
PC	Performance criterion
PDR	Policy Development and Review Department
PE	Public enterprise
PEM	Public expenditure management
PRGF	Poverty Reduction and Growth Facility
PRSP	Poverty Reduction Strategy Paper
ROSC	Report on the Observance of Standards and Codes
SAF	Structural Adjustment Facility
SBA	Stand-By Arrangement
TA	Technical assistance
VAT	Value-added tax
WEO	World Economic Outlook

Fiscal Adjustment in IMF-Supported Programs

Summary of Findings and Recommendations

This evaluation examines various aspects of fiscal adjustment in IMF-supported programs. This summary chapter sets out the framework that has guided the evaluation, explains the main findings and conclusions, and presents our recommendations for the future. It has been drafted in a self-contained manner.

Framework

Fiscal adjustment has traditionally been regarded as critical for achieving macroeconomic balance and is, therefore, often a central element in IMF-supported programs. It has also been the subject of much controversy on two grounds. The first relates to what may be called the *quantitative* dimension of fiscal adjustment, that is, whether, as some critics of the IMF contend, the fiscal component in programs reflects a "one-size-fits-all" approach often leading to excessive contraction. Such a contractionary bias can arise for two reasons:

- The programmed reduction in the external current account deficit may be larger than necessary in the sense that external resources to support a higher deficit could have been mobilized. This concern arises typically in low-income countries if the program design is unduly pessimistic about the prospects for concessional flows.

- Fiscal adjustment can also be excessive if programs are too optimistic in projecting recovery in the level of private demand, especially investment, during the adjustment process. In such situations actual private investment is much lower than projected and the fiscal adjustment programmed is therefore excessive. There is a case for fiscal policy playing a countercyclical role in such situations, though the scope for this depends upon other factors, such as the prospects for financing larger deficits and possible adverse market reactions to larger deficits because of debt sustainability.

The second set of issues which is potentially controversial may be called the *qualitative* dimension of fiscal adjustment. This relates to whether, given the scale and time path of fiscal deficit reduction, the efficiency, sustainability, and equity of fiscal adjustment could have been improved by using a different sequence and composition of policy measures on the revenue and expenditure sides. A core issue is how to match the short-term time frame of a program with the longer time frame often necessary to carry out the reforms, including institutional reforms, needed to create a more robust and resilient fiscal system able to withstand better shocks in the future.

The main data sources used in the evaluation are (1) a large cross-country sample of programs in the 1993–2001 period; and (2) more detailed desk studies of 15 specific IMF-supported programs, 4 of which were supplemented with analysis by local experts. The database used includes programs under the Enhanced Structural Adjustment Facility and the Poverty Reduction and Growth Facility (ESAF/PRGF), and Stand-By Arrangements/Extended Fund Facility Arrangements (SBA/EFF) in both transition and non-transition countries, some of which represent capital account crises during the period. These subgroups represent special categories and are recognized as such. Since the IEO has recently completed a report dealing with capital account crises,[1] and an evaluation of PRGF arrangements is currently under way, this evaluation focuses more on fiscal adjustment in SBA/EFF types of arrangements. An evaluation of IMF technical assistance (TA) is also part of the work program of the IEO for FY2004; consequently, the current project does not attempt to analyze the impact of TA in the fiscal area.

[1]See IEO (2003).

Findings: Quantitative Aspects of Fiscal Adjustment[2]

Are fiscal targets set on a "one-size-fits-all" basis?

The evidence does not support the view that IMF-supported programs adopt a one-size-fits-all approach to fiscal adjustment. The average targeted fiscal adjustment in 133 programs was 1.7 percent of GDP (1.4 percent for the primary balance) with a great deal of interprogram variation. The evidence also does not support the perception that programs always involve austerity by targeting reductions in current account and fiscal deficits or in public expenditures. In fact, in 40 percent of programs the current account deficit was projected to widen. Primary fiscal deficits were also programmed to widen and primary expenditures to increase as a percentage of GDP in slightly over one-third of cases.

In principle, the size of the fiscal adjustment proposed in each case should depend upon country-specific circumstances. They include the scale of the adjustment needed in the current account and the associated reduction in absorption to achieve this adjustment, market perceptions of the need for fiscal adjustment in view of debt sustainability problems, and allocative considerations relating to the balance between the public and private sectors. Cross-section analysis provides some insights on possible determinants of the targeted fiscal adjustment:

- The targeted adjustment seems to respond to both the initial fiscal deficit and the initial level of public expenditures. Countries with larger initial deficits and larger levels of expenditures in relation to GDP tend to have larger programmed deficit reductions.

- There is a significant positive association between the targeted fiscal adjustment and the envisaged change in the external current account. However, on average, only a small fraction (one-fifth) of the targeted change in net external financing is reflected in a corresponding change in the targeted fiscal deficit.

- The composition of the fiscal adjustment reflects initial levels of revenues and expenditures. Increases in revenues are programmed when initial revenues are low and reductions in expenditures are envisaged when initial expenditures are relatively high, and vice versa.

- In the ESAF/PRGF arrangements, two-thirds of the fiscal adjustment on average was programmed to come from the expenditure side. In contrast, in the SBA/EFF-supported programs in nontransition economies, two-thirds of the fiscal adjustment was targeted to come from the revenue side. In the transition economies, both revenues and expenditures were targeted to decline, reflecting the declining role of the state.

- On average, programs targeted a fiscal adjustment of about 1 percentage point of GDP across all types of arrangements during the first year of the program. This figure seems quite stable across different subgroups. Except for the transition economies, this represents between one-half and two-thirds of the total fiscal adjustment over a two-year period.[3]

- The different role that revenues and expenditure adjustments were expected to have during the lifetime of the program is particularly marked in the case of SBA/EFF in nontransition countries. In fact, spending as a share of GDP was not envisaged to decline but rather to increase in the first year, being offset by robust revenue performance to bring about a reduction in the fiscal deficit. The expected relative contributions of revenue and spending are sharply reversed during the second year of the program, when spending reductions become more important.

Surprisingly, the rationale for the proposed fiscal adjustment is not very clear when we look at the 15 individual programs studied in this evaluation. An in-depth examination of staff reports and other Executive Board papers related to these programs reveals that these documents often do not explain adequately how the magnitude and pace of the programmed fiscal adjustment have been determined. Nor do most documents explain how the fiscal targets relate to the rest of the program, in particular to assumptions about recovery in private sector demand and short-term growth prospects.

Did programs achieve their fiscal targets?

On average, programs achieved only about one-half of the programmed improvement in overall and primary fiscal balances. However, there is, once again, significant variation around this average. About 60 percent of programs underperformed but 40 percent overperformed with respect to pro-

[2]All macroeconomic magnitudes referred to here are in relation to GDP. All the changes are between the preprogram year and the second year after the start of the program.

[3]In the transition countries all the fiscal adjustment took place in the first year of the program. However, this was also the result of having a lower envisaged fiscal adjustment over a two-year period.

grammed deficit targets. The highest incidence of shortfalls was for SBA/EFF-supported programs in nontransition countries and the lowest was for SBA/EFF arrangements in transition countries.

Almost all fiscal adjustment on average takes place during the first year of the program. Except in the transition country arrangements, programs were unable to achieve further fiscal gains in the second year of the program in spite of more ambitious fiscal targets.

Cross-section analysis of the subset of programs that experienced shortfalls in fiscal performance suggests the following:

- Fiscal balances on average did not improve throughout the first two years of the arrangement—either in terms of overall or primary balances—except in the transition economies. Thus shortfalls appear to reflect weak fiscal performance rather than very ambitious fiscal targets.

- Overoptimism about fiscal adjustment is partly caused by overoptimism about growth projections. Absolute levels of revenue respond to growth with shortfalls in growth leading to corresponding shortfalls in revenue. However, absolute levels of expenditures, projected on the basis of optimistic growth forecasts, do not fall when growth falls below expectations, leading to an increase in expenditure ratios.

- There is a marked difference in the nature of fiscal shortfalls between programs that target a "large" fiscal adjustment (defined here as more than 3 percentage points of GDP over a two-year horizon, a definition that covers about 30 percent of the total sample) and others. In the latter group, excess expenditure as a share of GDP was the most frequent cause of the deficit shortfall, particularly in the nontransition countries.

- In contrast, revenue shortfalls were much more important in explaining shortfalls in performance in cases of "large" targeted fiscal adjustment. This pattern, which appears to hold both for concessional arrangements and programs supported by SBA/EFF arrangements, suggests that when substantial deficit reduction was judged necessary, programs aimed to achieve it through a combination of significant increases in revenues and cuts in expenditures.[4] However, in practice, the revenue increases achieved were much smaller, while the targeted expenditure reductions were generally achieved—perhaps forced by financing constraints.

The extent of expenditure adjustment appears to vary according to the initial fiscal imbalance. When initial fiscal deficits are moderate, expenditure as a share of GDP was little reduced if at all, notwithstanding programmed reductions. However, when the initial deficit was large, much of the fiscal adjustment was ultimately fulfilled through spending cuts. Expenditures seemed to be reduced only if strictly necessary and only if financing possibilities were unavailable. Efforts to increase revenues in situations of substantial fiscal imbalance generally fell well short of target; this pattern has important implications for structural reforms in the fiscal area, which are discussed later.

Flexibility of fiscal targets

IMF-supported programs are sometimes criticized on the grounds that they are insufficiently flexible, forcing a rigid pattern of fiscal adjustment that is not sensitive enough to changes in circumstances. The cross-country evidence does not support this view. A high proportion of the programs studied (about two-thirds) had incorporated revisions to their initial fiscal deficit targets by the completion of the second program review.[5] Of course, measuring the proportion of program targets that are revised is a rather narrow test of fiscal flexibility; it proves nothing about the appropriateness of any revisions. Nevertheless, it is important to note that in practice fiscal targets are revised frequently and these revisions are often associated with revisions in growth prospects. The cross-section data also suggest an interesting asymmetry in the process of revision: fiscal targets are revised downward when growth is below expectations, but they are less often revised upward when growth turns out to be higher than originally projected.

An examination of program and related documents suggests that the rationale for revisions is not clearly brought out. In particular, program documents often do not identify clearly what part of the fiscal shortfall was the result of exogenous developments (or unrealistic assumptions in the original program) and what part reflected a weaker policy effort. If, as often seems to be the case, insufficient progress in fiscal structural reforms is an important factor behind fiscal shortfalls, this needs to be frankly acknowledged in program reviews, and this is often not the case at present.

[4]As noted earlier, the pattern of fiscal adjustment in transition economies is somewhat different, since both revenues and expenditures are targeted to decline. However, in these cases also, revenue shortfalls are also large in programs that targeted a "large" deficit reduction.

[5]These are programs for which reviews are completed, that is, that remain "on track."

What has happened to economic recovery under programs?

A robust empirical investigation of the impact of IMF-supported programs on the pace of economic recovery is beyond the scope of this evaluation, and would involve comparing actual outcomes with the counterfactual of what would have happened to economic performance without a program or with an alternative program design. There is already a large, albeit inconclusive, literature on this topic.[6] Our analysis of actual and projected growth in a large sample of programs suggests the following conclusions:

- Average growth rates for different groups do not reveal a general tendency for growth rates to decline in program years, compared with the trend in the preceding decade. However, these averages mask considerable cross-country variation and growth did slow down, especially in the first program year, in a significant number of cases. The experience of the group of capital account crisis cases is particularly noteworthy since the average growth rate for this subgroup was negative in the first program year.

- While IMF-supported programs did not suffer from a generalized decline in growth, they did suffer from overoptimism. Except for the subgroup of transition countries (where the growth outcome was marginally better than programmed) average growth outcomes over a two-year horizon were lower than projected.

- Optimism regarding growth recovery was particularly significant in programs that started from an adverse situation. When growth was negative during the first year of the program, growth projections for the second year were on average twice as high as in reality. Moreover, programs were generally reluctant to project a slowdown in growth and very rarely projected negative growth. For example, growth slowdowns between the first and second year of the program occurred twice as often as they were projected.[7] Negative growth for the second year of the program was projected in only 1.3 percent of cases, but in reality it happened 10 times as frequently.

- Programs were also overoptimistic in projecting investment rates. Actual investment rates in the second year of the program were below projections in 60 percent of cases in a sample of 83 SBA/EFF arrangements. In about one-fourth of cases, investment rates were 5 percentage points of GDP or more below projections. Moreover, programs projected a decline in investment rates in one-fourth of cases while in reality investment rates declined in one-half of the arrangements.

Growth optimism, and especially the reluctance to forecast downturns in programs, has many causes, including especially the understandable desire of both the IMF and the authorities to present a relatively upbeat recovery scenario. However, this has important implications for program design because it understates potential risks and preempts a systematic discussion of the appropriate role of fiscal policy in the event of a significant economic downturn. This was clearly a major factor in the capital account crisis cases in East Asia, where—as suggested by the recent IEO study of three capital account crisis cases—adverse balance sheet effects on private demand were underestimated.[8] It also seems to have been a factor in many other SBA/EFF-supported programs in nontransition economies.

Is there a contractionary bias in fiscal design?

The fact that both output and investment appear to be consistently lower than projected raises the issue whether there is a contractionary bias in fiscal design. Critics have argued that IMF-supported programs would ensure quicker recovery if they anticipated weak investment demand more accurately and therefore adopted a less contractionary stance of fiscal policy. A tight fiscal stance is not inappropriate when it is assumed that private investment demand is buoyant and fiscal contraction creates room for private investment to be financed. However it is not appropriate in situations where there is a sharp downward shift in the investment function, or when the level of private demand responds much more sluggishly to the program than originally projected. There is evidence that investment is consistently overestimated in IMF-supported programs and there is overcorrection of the current account deficit. In a large number of the cases the overperformance in the current account deficit is combined with an excess buildup of reserves, suggesting that the economy could respond positively to a demand stimulus. In such situations, it could be argued that a less contractionary fiscal stance might have been appropriate.

This conclusion needs to be qualified in one important respect. It focuses only on the role of fiscal

[6]A review of the literature on this topic can be found in Haque and Khan (1998).

[7]Programs tend to underpredict significantly more situations of adverse output developments than underpredict situations of favorable output developments.

[8]IEO (2003).

adjustment via its impact on aggregate demand. However, emerging market countries relying upon international financial markets also have to consider the impact of their fiscal stance on market confidence and the resulting availability of external finance. Where debt sustainability is an issue, it may be desirable to adopt a tighter fiscal stance than justifiable on countercyclical grounds alone to ensure a quicker return to confidence.

It is difficult to determine in any particular case how to weigh these different considerations and come up with a fiscal stance that provides an appropriate balance. However, these issues need to be explicitly discussed and explained in program documentation. One of the conclusions of our evaluation is that this is not done in a systematic way. Board documents generally provide insufficient analysis and justification for the proposed fiscal adjustment path or the assumptions driving the projected recovery of private spending and how it is linked to program instruments, including the fiscal stance. Inclusion of such an analysis would help to avoid growth overoptimism. It would also provide a more coherent framework for sensitivity analysis that would alert staff early on in the process to what should be monitored as the program unfolds. We recognize that fiscal fine-tuning to take account of all these factors is extremely difficult and, in practice, a large part of the outcome must be based on judgment. However, a more explicit discussion of the key macroeconomic assumptions underlying the proposed fiscal path would promote greater understanding of the risks and uncertainties involved and also facilitate necessary mid-course corrections in the fiscal stance. Many such mid-course corrections do occur in practice, but their rationale is often unclear. A clearer statement of the original rationale would permit a more transparent basis for adjusting fiscal targets in the course of program implementation.

Internal review process

An examination of the internal review process, focusing on the comments of the Policy Development and Review Department (PDR) and Fiscal Affairs Department (FAD) on the fiscal aspects of the 15 individual programs studied in this evaluation suggests the following:

- Internal review comments do pay attention to the need to justify the specific fiscal stance, but (as noted above) these comments do not lead to an explicit analysis in the final Board documents of the factors that led to the determination of the fiscal stance. The possibility that projections of private sector activity and growth recovery were overoptimistic was generally not given much attention in the review process.
- In many cases, the scope and detail of review department comments was greater at the stage of program reviews than at the stage of initial program design. A comprehensive internal debate would have the greatest value added if it took place at an early stage of program formulation and involved an exploration of alternative policy options to achieve broad objectives. This approach would also be more conducive to encouraging domestic ownership of programs. Instead, the review process is much more reactive, with reviewers commenting increasingly as programs proceed, instead of at the design stage. This may reflect relatively sanguine initial judgments (associated with overoptimism in growth prospects, and policy implementation) that the fiscal and other targets would be achieved, followed by a closer look as revisions become necessary. We understand from staff that there are often considerable informal consultations on key design issues before the formal briefing paper stage. However, these are not substitutes for a more active examination of risks and options in the initial stages. The fact that Board documents in the programs we examined incorporated overly optimistic assumptions, and did not specify the links between the fiscal stance and the recovery of private activity and output, should be a matter of concern.

Social Spending and Social Protection in IMF-Supported Programs

The impact of IMF-supported programs on the level of public spending in the social sectors has received a great deal of attention, with many critics voicing concern that these programs typically involve an unnecessary squeeze on social expenditures. The evaluation examines this issue in several ways.

Projections of aid flows in concessional programs

Concerns have been raised that IMF-supported programs in low-income countries (that depend on concessional financing) may incorporate fiscal targets based on aid projections that "taper out" too quickly relative to what donors may be willing to provide. Some have suggested that this feature of program projections may in itself create a disincentive for donors to sustain their level of aid—even when programs remain on track.

To address this issue we have examined program projections for nearly 100 ESAF/PRGF programs approved in the period 1995–2001, complemented

by an in-depth study of a sample of 20 concessional programs in sub-Saharan African countries. The results show that program projections of aid do tend to decline over the medium term, albeit at a moderate pace in most cases. However, there is no evidence that projections have systematically underestimated actual aid flows for the outer years of programs. The analysis used here cannot answer the much more complex question, which goes beyond the scope of the present evaluation, of whether more ambitious program targets for public expenditure (and deficits) could have resulted in the mobilization of additional concessional financing from donors.

Effect of IMF-supported programs on the level of social spending

There has been a long-standing debate on the impact of IMF-supported programs on public sector social spending. We address this issue through an econometric analysis of 146 countries from 1985 to 2000, looking at years with and without an IMF-supported program. In order to assess the impact of programs on expenditures in health and education, we controlled for other factors affecting social spending as well as for the endogeneity of the presence of an IMF-supported program.

The results show that the presence of an IMF-supported program does not reduce public spending in either health or education—measured as a share of total public spending, GDP, or in per capita real terms. In fact, we estimate that during program periods, and with all other factors being the same, public spending in each of the health and education sectors increased by about 0.3 to 0.4 percentage points of GDP compared to a situation without a program. This increase is sustained beyond the end of the program but it diminishes over time.

From the fact that social spending increases, it is not possible to argue that the most vulnerable groups of the population are effectively protected from the economic shocks they may experience during program years. This will depend on how that increased spending is targeted and timed. Unless governments already have in place programs and budgetary mechanisms that allow for that protection, IMF-supported programs generally have too short a time frame and the IMF lacks the necessary expertise to assist in implementing such policies. This suggests that an alternative framework may be needed to address such issues.

Social concerns in program design

Current practices of the IMF in the area of social protection in non-PRGF countries follow the 1997 Guidelines on Social Expenditures, which call for the IMF staff to track health and education spending and, by relying on work by the World Bank, encourage authorities to incorporate spending targets for these sectors in the Letters of Intent (LOIs) that spell out program objectives. The guidelines also encourage staff to monitor trends in basic social indicators (such as infant mortality and school enrollment) drawing on the World Bank. However, the guidelines are quite broad and general in scope, and discussions with staff suggest that there is considerable uncertainty about what is expected in practice, at least outside the PRGF/PRSP countries. There also appears to be some uncertainty among the staff as to how the initiative to streamline conditionality should affect the IMF's approach in this area.

A detailed examination of the 15 sample programs (complemented by 8 additional more contemporary programs to gauge recent progress) shows substantial variation in how social expenditure issues are treated in practice. Trends are noted in some program documents for broad categories of expenditures such as education and health. However, only one-third of the sample of 15 programs analyzes these trends and identifies priority social expenditures that need protection—although the most recent group of programs shows limited improvement in this respect. Performance criteria were rarely used to support social measures; however, 9 of these 23 programs used benchmarks or indicative targets. Only half of the more recent 8 program documents analyze changes in social spending and few programs (outside the PRSP/PRGF countries) discuss how explicit monitoring and feedback systems could be established or how these aspects would be integrated with the work program of the World Bank. Thus, the empirical basis for establishing and assessing policy actions in this area is often absent.

The internal review process by PDR and FAD quite often gave feedback in this area—providing specific suggestions to design and support priority social programs to protect vulnerable groups. Most of these comments, however, were concentrated in the review phase during program implementation, and hence were too late to influence basic program design.

An important finding from the case studies is that it is not necessarily costly to preserve critical programs or budgetary allocations to protect the most vulnerable groups from external shocks or budgetary retrenchments. This can be facilitated by some reallocations in the budget—a possibility particularly relevant for middle-income countries. However, the objective of protecting critical expenditures cannot be achieved simply by monitoring trends in broad social spending categories. Such monitoring would likely fail to capture micro-level reallocations that tend to take place in periods of fiscal stress that undermine social protection. As discussed in this report, spend-

ing categories that often are most critical to vulnerable groups come under pressure and are likely to be preempted by other expenditures during these periods (e.g., basic medical or primary school supplies being preempted by personnel expenditures).

The protection of critical spending categories and well-targeted programs in the social sector can thus play an important role in protecting the most vulnerable from adverse shocks and budgetary retrenchments at fairly low cost. Efforts should, therefore, be made to build such elements into program design whenever possible. This emphasis is consistent with the IMF Articles of Agreement, especially Article I (v), which states that one of the purposes of the IMF is to make "the general resources of the Fund temporarily available to [members] . . . providing them with opportunity to correct maladjustments in their balance of payments without resorting to measures destructive of national or international prosperity." It would also help to make more concrete commitments by the IMF to "minimize the adverse effects [of macroeconomic adjustment on vulnerable groups] and, when some are inevitable to achieve the desired reforms, to mitigate these effects through compensating measures."[9]

To be effective in this area, the IMF would need to work within an operational framework that takes account of four constraints: (1) policies in this area must be truly homegrown and fully owned; the major initiatives must therefore come from the country; (2) since the IMF does not have expertise on social sector issues, nor is this an area of its comparative advantage, inputs from other agencies, especially the World Bank, are critical; (3) there is a mismatch of time frames between the short-term nature of IMF programs and the longer-term time frame needed for building institutions and budgetary systems that can provide social support in times of crisis effectively; and (4) finally, it is necessary to ensure that the attempt to incorporate social protection into IMF programs does not contradict the recent streamlining initiative by leading to an overload of conditionality.

In the case of low-income countries, the PRSP framework is expected, in principle, to meet these requirements. However, there is at present no framework for non-PRGF eligible, predominantly middle-income, countries that would ensure identification of critical and homegrown social sector support programs that could be used as mechanisms for social protection at the time of crisis. The PRSP framework is obviously not appropriate for middle-income countries, but in the absence of any framework there will be a growing divergence between the way these critical social issues are treated between PRGF and non-PRGF countries. It is, therefore, necessary to revisit the 1997 guidelines with special reference to what IMF staff should do consistent with the new emphasis and special constraints discussed above.

Some elements of a workable approach can be readily identified. First, the mismatch of time frames suggests that necessary preparatory work in this area must be undertaken not at the time of crisis but much earlier, as part of normal surveillance. In order to ensure that initiatives are homegrown, the IMF could request governments to consider identifying critical social spending to be protected, or safety nets to be activated, in the event of crisis. The IMF could also encourage countries to approach the World Bank for assistance in this area. The IMF on its part, consistent with its mandate, could report on the authorities' responses in this area and monitor progress.

Building on recent initiatives (such as the call for increased coordination on public expenditure management (PEM) issues), both institutions could work to develop a broad understanding with the authorities on the reforms needed and an appropriate sequencing for implementation. Where joint efforts are required, for example, in public expenditure management, a country-led work program would be jointly established. On the basis of the resulting joint effort, the IMF and the World Bank could assist the authorities in setting up mechanisms to track critical social spending through the budget and identify ultimate allocations, including to local governments where a significant amount of spending is decentralized.

Reforms in the Fiscal Area Under IMF-Supported Programs

An important part of the shortfall in fiscal adjustment results from optimism regarding the pace of implementation of structural reform on the fiscal side. Moreover, much of the fiscal adjustment achieved is through measures that do not assure long-term sustainability and flexibility of fiscal systems to future shocks. We have looked at three dimensions of reform policies in the fiscal area: (1) the balance among various policy measures, whether programs tilt toward specific areas while neglecting others; (2) the progress in implementation; and (3) the role of surveillance in helping the process of reform.

Balance among policy measures emphasized by programs

Fiscal adjustments in programs have focused more on the revenue side than on reallocations and reforms on the expenditure side. On the revenue

[9]IMF (2000a).

side, the accent has been on introducing or increasing value-added tax (VAT) rates, with less attention paid to income and property taxes and tax administration efforts aimed at reducing evasion. Sometimes these VAT rate increases have been resisted by broad segments of the population because they have been perceived to be inequitable relative to other revenue-raising possibilities.

The VAT needs to continue being promoted as the cornerstone of a modern tax system. However, stronger and parallel efforts should be made at improving collections, curtailing discretionary exemptions, and reducing tax evasion—particularly direct taxes (personal and corporate) and customs duties. Even in the short run, these efforts could yield important revenue increases if targeted at collecting from well-known taxpayers with arrears or those believed to be significantly underpaying (hence reducing the need for large increases in VAT rates to quickly generate revenues). When tax authorities have displayed determination in this area, the results have been impressive and have received wide support. This evaluation finds that efforts by the IMF in this area have not been forceful enough, both in the context of programs and in surveillance, particularly if they affect powerful vested interests. Often, tax administration reforms in IMF-supported programs have focused on the technology side rather than on politically more difficult actions, such as legislation to empower tax agencies to pursue tax evasion forcefully and for the system to be less prone to political interference.[10] More forceful actions in this area may also increase the support of society at large for the overall reform agenda supported by programs.

Improving tax collection and reducing exemptions and evasion is an aspect of fiscal reform that should be pursued more vigorously. Estimates and comparisons of the extent of tax evasion should be made public, drawing where possible on cross-country analysis. These steps require both political will and institutional changes, in different mixes according to the specific situation, and should be unbundled.

On the expenditure side, an examination of the different programs shows that conditionality has been concentrated on short-term quantitative targets to reduce public employment or cap public sector wage increases (which generally prove to be short lived because they are easy to reverse) rather than focusing on the reorientation of public spending and medium-term civil service reform. As a result, progress in reducing the wage bill has been neither sustainable nor efficient—reversals have often occurred.

The internal review process often addresses these areas of weakness, including the need for expenditure reallocations, and (perhaps most important) the need for determined actions by the executive in the areas of reducing tax exemptions, limiting tax incentives, and taking concrete actions against tax evasion and tax arrears. But again, as in other areas, these comments come too late in the process to influence initial program design.

Progress in implementation

Our evaluation shows that progress in implementing fiscal reform initiatives in the sample of 15 programs was limited. In no given reform area was implementation satisfactory in more than 40 percent of cases. Measures to reduce the public sector wage bill, achieve civil service reform, and reform the social security system have been particularly difficult to implement.

This limited progress is often the result of an excessive emphasis on measures to meet short-term quantitative targets, rather than a focus on critical institutional changes that might extend beyond the end of the program. This is largely the result of a mismatch of time frames, such as the short horizon of programs relative to the time needed to complete these institutional reforms. Such reforms may need to be broken down into several steps: some of them can be started at the outset of the program with enough determination from the executive branch; others will require time to the extent they call for legislation and improvements in the implementation capacity of agencies. Surveillance could play a key role in providing such a road map, but, as the next section suggests, it often does not do so.

Learning and the process of surveillance

We have examined program request documents to assess the extent to which they look at the past in order to draw lessons. We also examined surveillance activity over the three-year period prior to the IMF lending arrangement in the 15 sample programs studied.

Program requests are only partly successful in evaluating past fiscal performance (with an index of success of 50 percent).[11] The results are worse

[10]As documented in Appendix 7, the IMF has provided extensive technical assistance (TA) in this area. Since the focus of this evaluation is on fiscal adjustment in IMF-supported programs, we have not examined IMF TA here. Our findings here should not be interpreted as indicating failures in technical assistance, which is clearly targeted at addressing the technology of fiscal reform. Our concern is whether programs have been successful in encouraging politically difficult decisions regarding tax collection, decisions that are critical to take advantage of the technical solutions proposed by TA.

[11]The index, discussed in Chapter 7, is based on assigning weights to programs with good, mixed, and poor performance. Thus, it has inevitably a measure of subjective judgment.

(35 percent success) when documents are judged on whether they analyze policy failures under the prior arrangement. Overall, programs tend to focus on fiscal performance during the last year prior to the program, and rather independently of previous arrangements. Few efforts are made to analyze the factors behind policy failures.

We have also examined the link between surveillance and programs. Although there is significant variability, efforts during surveillance to forcefully flag the need to accelerate reform (in areas where implementation was lacking) have been limited, with an index of success of 40 percent. Surveillance is drawing too few lessons from past failures, often not setting future paths for more complex reforms.

Focusing on the unfinished reform agenda will require strong follow-up during surveillance, as well as continuity in successive programs. Our results suggest that surveillance does not forcefully flag policy inaction—many times it is insufficiently candid in language. Although based on a very small sample, self-standing surveillance does not seem to yield better results. This is a missed opportunity because we would expect that surveillance not associated with a program request or review would have a genuine opportunity to take a more strategic perspective on both, whether fiscal reforms over time add cumulatively to better fiscal systems, and what the remaining fiscal agenda for the future should be.

Surveillance could play a much more forceful role in providing a medium-term road map of structural reforms to be followed up over time, with or without programs. Progress and reasons for inaction should be reported candidly. That road map could then provide guidance for the specific reform priorities to be taken up in successive programs—this being particularly important in repeat users of IMF resources.[12]

Recommendations

Based on the conclusions of our evaluation of fiscal adjustment in IMF-supported programs we propose five recommendations for the consideration of the Executive Board that in our view would help overcome the weaknesses identified in the evaluation**.**

Recommendation 1. Program documentation should provide a more in-depth and coherent justification for the magnitude and pace of the fiscal adjustment and how it is linked with assumptions about the recovery of private sector activity and growth. The evaluation shows that, while the criticism of a one-size-fits-all approach to fiscal adjustment in programs is not correct, the rationale of the fiscal adjustment projected (in terms of the various possible factors that are relevant) is not adequately spelled out. There is also a tendency to underestimate the potential for economic downturns and/or to be optimistic about output recovery. Because these optimistic forecasts are usually predicted on the behavior of private demand, the relative role of the fiscal stance in either complementing that demand or releasing resources to finance it may be subject to some systematic bias. A more explicit articulation of the basis for the proposed fiscal stance, and how it is linked with assumptions regarding the recovery of the private sector, will help to promote a better understanding of the various factors involved and also to identify possible risks and subsequent corrective measures. It will also facilitate the review process and discussions at the Board, as well as provide external audiences with a more convincing explanation for the rationale for the program.

Recommendation 2. The internal review mechanism should place relatively more emphasis on the early stages of the process. Our evaluation shows that reviewers raised many questions, and also provided rich inputs into areas identified as relatively weak in this evaluation, but most of them came late, when there was little scope for effective program design. A more intensive process of brainstorming is needed at the time of the initial brief, and that brief should also articulate more clearly the basis for the fiscal program, including debt sustainability issues.

Recommendation 3. Programs should give greater emphasis to the formulation and implementation of key institutional reforms in the fiscal area, even if (as is likely) they cannot be fully implemented during the program period. The evaluation results suggest that slow progress in implementing structural and institutional reforms in the past puts limits to the quantity of fiscal adjustment that can be achieved by a program in the short run. A greater emphasis on structural reforms relative to the establishment of detailed quantitative targets will ultimately enhance the ability of fiscal systems to achieve more durable adjustments and handle shocks in the future.

In making this recommendation, we are not suggesting abandoning short-term quantitative targets, nor do we believe that the proposed greater emphasis on structural reforms will reduce the need for fiscal adjustment in the short term. Short-term adjustment is often unavoidable in a crisis and firm action is needed in such cases. However, programs should make much stronger efforts to specify those structural reforms which should be carried out during the

[12]The IMF's Executive Board has already indicated that a more forward-looking strategic assessment is required in cases involving prolonged use of Fund resources and that surveillance could be a suitable vehicle for reporting such assessments.

program horizon as part of a broader road map of priority reforms. This road map, and its prioritization, should ideally have emerged in the course of surveillance and be updated regularly as outlined in Recommendation 4 below.

Recommendation 4. The surveillance process should be used more explicitly to provide a longer-term road map for fiscal reforms and to assess progress achieved. The evaluation finds that a significant constraint in improving the quality and sustainability of the fiscal adjustment is the mismatch of time frames between the short horizon of the typical IMF-supported program and the longer time frame required to implement most structural reform measures. Programs in PRGF countries already have a framework (the PRSP) for longer-term IMF involvement that allows the follow-up of such reforms. This is not the case for arrangements in non-PRGF countries. Thus, an increased divergence in approaches is emerging between these two sets of countries. To some extent, this is inevitable: indeed, it reflects an explicit decision by the international community to adopt a different approach to adjustment in the low-income countries, in recognizing their particular needs. However, many non-PRGF countries also face important second-generation fiscal reforms that require significant time. While the operational framework for addressing the mismatch of time frames will inevitably be less precise in these cases—especially if the IMF's program involvement is relatively infrequent—the fiscal aspects of surveillance could be strengthened to provide such a framework.

We recommend the following specific steps:[13]

- In collaboration with the authorities, the IMF should clearly identify in surveillance reports the most critical distortions in a country's public finances from the perspectives of equity and efficiency. These distortions could be summarized, for example, in the form of a box or matrix analyzing key "Fiscal Reform Priorities."
- Such an analysis would provide a road map for fiscal reform in the future, with a clear sense of priorities. It would help to provide the basis for identifying critical reforms—particularly in areas where these reforms have been lagging—that would need to be addressed should IMF financing be required in the future.
- The identification in advance of areas considered critical is not intended to predetermine future conditionality in a mechanical fashion. Rather, it will allow the authorities flexibility in the timing and packaging of reforms which is often lost if these reforms are flagged at the last minute in the context of a crisis situation. This approach would also help foster greater domestic debate on key reforms and hence would encourage homegrown solutions and greater ownership. Early and clear prioritization of reforms is also consistent with streamlining objectives—it will avoid last-minute bunching of reforms under crisis situations.
- The analysis of fiscal reform priorities should be accompanied by an assessment of why certain important distortions were not addressed in the past and what are the lessons from past experience. This should include an effort to identify and unbundle the various constraints to critical reforms, including lack of technical capacity, areas where additional legislative action is necessary, and areas where key decisions from the executive branch are required.
- Work in the fiscal area in the course of surveillance should include more systematic efforts to estimate the extent of tax evasion and tax exemptions, including the use of cross-country comparisons.
- Public debt sustainability analysis is now increasingly being carried out following the recent Board paper on sustainability.[14] This work could help anchor the road map of fiscal reform priorities proposed above and to assess trade-offs over time. At the same time, debt analysis provides a check of cumulative progress in improving fiscal systems that could also be reported in successive surveillance reports.

We recognize that there are many priorities for surveillance, and some selectivity will be required. The attention to be devoted to those issues need not be the same in all countries. One approach could be to use surveillance to identify countries where the fiscal situation is especially stressful and to conduct an in-depth fiscal surveillance exercise with such countries, with subsequent updates every three to four years. There would be considerable merit in coordinating such exercises with the work program of the World Bank.[15]

[13]While the focus of this evaluation is on IMF-supported programs, the recommendations discussed here—which aim to strengthen the fiscal aspects of surveillance and give it a longer-term perspective—are relevant for all IMF member countries.

[14]IMF (2002e) and more recently IMF (2003b).

[15]Such collaboration is consistent with the proposals for "systematic information sharing and monitoring in the context of lending operations and in CAS and Article IV consultations." IMF and World Bank (2002). The in-depth fiscal surveillance exercise could also be helpful in identifying cases to conduct a fiscal management assessment (FMA) along the lines of the Turkey FMA. Turkey was the first country to benefit from an FMA (SM/02/191, 6/20/2002), which assessed and provided suggestions on how to improve the transparency and coordination of institutions in the area of fiscal policy.

Recommendation 5. The IMF should clearly delineate the operational framework in which social issues will be addressed within program design in non-PRGF countries. This should include a clear indication of the IMF's responsibilities and activities in this area. The present (1997) guidelines that direct IMF work in the social area remain vague and difficult to translate into operational policy advice. Evidence from the evaluation suggests that the result in practice has been a wide variation in approaches and a tendency to promise (in terms of general policy statements in program documents) more than can be delivered.

The objective should be to assist middle-income countries to prepare and improve their institutional framework to allocate resources to critical social programs and to establish mechanisms to protect the most vulnerable groups in the face of external shocks and budgetary retrenchment. Following the principles outlined in the section "Social concerns in program design," the framework could include the following elements.

- The IMF could invite the authorities regularly during Article IV consultations to suggest what are the existing critical social programs and social services they would like to see protected in the event of adverse shocks. Participation on the part of the authorities would clearly be voluntary.
- Successful implementation of efforts to protect expenditure on critical social programs and deploy social safety nets will depend heavily on having better and more transparent expenditure monitoring systems. On the basis of the priorities identified by the authorities, the World Bank and the IMF could agree with them on an accelerated work program on public expenditure management (PEM) systems, specifically geared toward the social area so as to protect the specified programs and spending categories. This provides the opportunity for a concrete application of the recent initiative discussed at the Board to increase coordination between the World Bank and the IMF in enhancing PEM systems.
- This concrete application of the PEM initiative is particularly important because in many cases where there is an IMF-supported program the World Bank is also active with adjustment lending supporting the budget. Joint work programs on PEM systems provide an ideal opportunity for both institutions to play an enhanced role in assisting in the protection of critical social expenditures during these periods.
- Surveillance would routinely report on these initiatives and their progress over time.

Chapter I Introduction

Fiscal policy is central to macroeconomic management and is, therefore, the subject of considerable attention in the course of Article IV surveillance and in the design of IMF-supported programs. In fact, it is often the centerpiece of program design, with quantified targets included as key elements of conditionality. Fiscal adjustment is also among the most controversial elements in IMF-supported programs. Critics complain that the scale of the adjustment is often unduly harsh and likely to impart a contractionary impulse at a time when economic activity is depressed in any case, thereby leading to unnecessary loss of output and employment, with adverse effects on the poor.[1] Apart from aggregate output and employment effects, fiscal adjustment is also controversial because of its potential distributional effects. Policies for reducing or constraining spending to meet fiscal targets are often criticized on the grounds that they squeeze socially beneficial spending such as health and education or withdraw subsidies on items of essential consumption, thus placing a disproportionate burden of the adjustment on those least able to bear it. Efforts to mobilize higher tax revenue are also sometimes criticized because many of the tax measures which can be introduced in practice in the short term, such as an increase in the rate of general sales taxes or VAT, are viewed as regressive.

This evaluation aims at examining the experience with fiscal adjustment in IMF-supported programs to shed light on these issues and make recommendations for surveillance and program design in the future.

The evaluation is based on analyses at two levels. Part of the evaluation relies on cross-section analysis using two large samples: the Monitoring of Fund Arrangements (MONA) database, which provides information on both program targets and actual outcomes for 169 programs approved during 1993–2001, and a database obtained from the Fiscal Affairs Department (FAD), which provides information on health and education spending covering 146 countries over the period 1985–2000. The cross-section analysis is supplemented by a more detailed examination of two smaller samples. We have examined program and associated surveillance documents for 15 programs in a mixture of low-income, transition, and middle-income countries to evaluate the internal mechanisms and processes through which fiscal targets are set and program performance reviewed. In 4 programs, we complemented the information with work by local experts.[2] We also examined 20 programs in sub-Saharan Africa to consider the specific issue of whether aid availability projections are unduly pessimistic, forcing an unnecessary contractionary stance in programs.

The databases used cover a variety of programs including ESAF and PRGF arrangements and SBA/EFF arrangements in both transition and nontransition countries, some of which represent programs in the context of capital account crises. These categories are separately identified in the analysis where necessary. Since the IEO has recently completed a report dealing with the role of the IMF in capital account crises[3] and a detailed study of PRGF countries is currently under way, this evaluation focuses less on these cases and more on the implications for fiscal adjustment in middle-income countries. This report is organized as follows:

- Chapter 2 examines patterns in the way IMF-supported programs set fiscal targets.
- Chapter 3 examines how well Board documents explain the rationale for such targets and their

[1]See for example, Center of Concern (1998); European Network on Debt and Development (2001); International Financial Institution Advisory Commission (2000); Oxfam (1995, 2001a, and 2001b); World Development Movement (2000a and 2000b); Watkins (1999); Kanbur (2000); Collier (1999); Collier and Gunning (1999); Stiglitz and Furman (1998); and Feldstein (2002).

[2]The programs included are Algeria SBA 1994, Bulgaria EFF 1998, Costa Rica SBA 1995, Ecuador SBA 2000, Egypt SBA 1996, Jordan EFF 1999, Pakistan SBA 2000, Peru EFF 1996, Philippines SBA 1998, Romania SBA 1999, Senegal PRGF 1998, Tanzania ESAF/PRGF 1996, Ukraine EFF 1998, Uruguay SBA 2000, and Venezuela SBA 1996. It was complemented by work by local experts in the case of Ecuador, Philippines, Romania, and Tanzania.

[3]See IEO (2003).

links to the rest of the program. The internal review process at different stages of the program is also examined.

- Chapter 4 analyzes actual fiscal performance and compares it with targets. It looks at the sources of shortfalls and how program reviews under implementation revise fiscal targets.
- Chapter 5 examines the experience with economic recovery under programs and the degree of optimism in program projections. It then looks at possible sources of contractionary bias in typical SBA/EFF arrangements.
- Chapter 6 examines several concerns regarding social spending in IMF-supported programs: whether there is a downward bias in projecting donor aid that may compress social spending, what has been the impact of programs on social spending, and how programs are taking into account social issues in program design.
- Chapter 7 analyzes the process of reform in the fiscal area in the 15 sample programs, including progress in implementation and how well surveillance is supporting the process.

CHAPTER 2

Fiscal Targeting in IMF-Supported Programs: Cross-Country Analysis

A common criticism of fiscal adjustment in an IMF-supported program is that it is derived from a "one-size-fits-all" approach, which places too much emphasis on fiscal adjustment (i.e., a reduction in the fiscal deficit defined in terms of either the overall deficit or the primary deficit) without taking account of the specific circumstances of the country. In this chapter, we examine the available evidence on fiscal targets in IMF-supported programs and the extent to which they vary across countries. First, we outline some of the considerations that should ideally be taken into account in setting fiscal targets. Next, we use cross-country analysis to examine a large number of past IMF-supported programs and assess patterns and statistical regularities in the way fiscal targets actually are set.

Relevant Considerations in Determining the Fiscal Stance

It is not easy to determine what should be the extent of fiscal adjustment in a particular country situation. There are several factors that are potentially relevant in determining the nature of fiscal adjustment and some of them could point in different directions.

(1) The scale of fiscal adjustment needed can be viewed as a function of the scale of adjustment required in the current account. Any given reduction in the current account deficit requires a reduction in domestic absorption, and a lower fiscal deficit is a way of reducing excess absorption in the public sector. This is the traditional reason for advocating contractionary fiscal policies in a situation where a reduction in the balance of payments deficit is needed. The need for fiscal adjustment is particularly evident when the current account deficit is bloated by fiscal expansion to begin with, since the alternative would be to force the private sector to bear the burden of adjustment which may fall disproportionately on investment.

(2) The fiscal deficit may need to be reduced as part of an adjustment program where concerns about the sustainability of public debt are expected to have negative effects upon capital inflows. This consideration is particularly important in emerging market economies that have achieved a degree of integration with international financial markets and that rely on financial flows that are highly sensitive to market perceptions regarding debt sustainability. The need for fiscal adjustment in such cases is driven not so much because of the necessity to reduce aggregate demand but rather by the need to persuade markets about debt sustainability to ensure a sufficient flow of resources to finance the existing current account deficit. The scale of the adjustment needed depends upon the stock of public debt in relation to GDP; the potential rate of growth of the economy; and also psychological factors, which determine market perceptions of growth potential and sustainability.

(3) It is also possible to envisage a reduction in the fiscal deficit driven mainly by allocative concerns: that is the desire to reduce the degree to which the fiscal deficit crowds out the private sector. The preexisting size of the fiscal deficit is clearly a relevant factor in determining the direction and scale of adjustment. The volume of government activity in relation to GDP is also important since high levels of government spending clearly signal that some crowding out has taken place.

The importance of these factors would obviously vary from country to country and one would, therefore, expect that the fiscal deficit target built into program projections reflects country-specific judgments on the importance of each of these elements. In the rest of this chapter, we use cross-section data to throw light on these issues, followed by an in-depth study of 13 programs in Chapter 3.

Fiscal adjustment envisaged in programs

Table 2.1 presents the average initial conditions for different types of programs in period *T*–1, the year immediately preceding the first program year. It

Table 2.1. Initial Conditions as Seen by Staff at the Start of the Program[1]
(In percent of GDP)

			SBA/EFF	
	All Arrangements	ESAF/PRGF[2]	Transition countries	Nontransition countries
External current account balance	–5.2	–7.0	–4.1	–3.9
Overall government balance	–4.1	–4.5	–4.2	–3.6
Government primary balance	–0.6	–1.0	–1.4	0.3
Government revenues and grants	24.4	21.4	33.4	22.9
Government expenditures	28.5	25.9	37.6	26.5
Growth trend (percent)[3]	1.6	1.8	–2.0	3.3
Annual inflation (percent)	92.3	35.7	355.3	14.1
Count (number of programs)	169	71	34	64

Source: Calculated from MONA database.
[1]Initial conditions are measured by outturns for the year immediately preceding the first program year (i.e., year *T*–1), as reported in the MONA database.
[2]Includes all arrangements under concessional facilities—SAF, ESAF, PRGF—including those that were combined with SBAs and EFF arrangements.
[3]For each arrangement, the average rate of real GDP growth in the 10 years preceding the initital program year.

Table 2.2. Program Projections: Changes in Balances from (*T*–1) to (*T*+1)
(In percent of GDP)[1]

			SBA/EFF	
	All Arrangements	ESAF/PRGF	Transition countries	Nontransition countries
Current account	–0.3	0.1	–2.0	–0.2
Government balance	1.7	1.6	1.1	2.0
Primary balance	1.4	1.0	0.4	2.0
Government revenue	0.4	0.4	–1.7	1.3
Government spending	–1.2	–1.2	–2.8	–0.7
Count[2]	133	60	21	52

Source: MONA database.
[1]Figures subject to rounding errors. Magnitudes over 0.5 percent of GDP are statistically significant (different from zero) except current account and government balances in transition countries.
[2]The sample size of 169 arrangements reported in Table 2.1 fell to 133 because data for the second year of the program were unavailable. Most of the reduction in sample size from 169 to 133 was accounted for by arrangements approved in 2001 for which no actuals were available for the second program year.

provides a background against which to compare the fiscal adjustment envisaged in programs.

Table 2.2 provides the average magnitude of envisaged change in fiscal and external balances in the original program design, for the sample as a whole and for the individual subgroups. Since a significant proportion of arrangements were approved well into the initial program year—nearly 40 percent were approved in the second half of the year—we examine changes in key variables over a two-year horizon from the year immediately preceding the initial program year (year *T*–1) to the end of the second program year (year *T*+1).

The following features of the projected changes presented in Table 2.2 are of interest:

- IMF-supported programs have, on average, envisaged only very small changes in external balances between *T*–1 and *T*+1. The only large change envisaged is in the case of transition countries where the current account deficit was projected to widen by 2 percentage points of GDP on average.

- The average targeted improvement in the fiscal balance for the sample as a whole is relatively modest, about 1.7 percentage points of GDP over two years. The programmed improvement in primary balances was even lower—1.4 percent of GDP—implying a slight reduction in envisaged interest payments as a share of GDP.

- The composition of the targeted fiscal adjustment shows, on average, a much larger reliance upon spending reductions than on revenue increases. This is true of the ESAF/PRGF group

Table 2.3. The Direction of Change in Selected Macroeconomic Targets in IMF-Supported Programs as a Share of GDP

(Initial level of balances are shown in italics, as percent of GDP)[1]

Panel A. Distribution of Programs According to the Direction of Envisaged Changes in Current Account and Government Balances[2]

		Current Account Balance		
		Deterioration	Improvement	
Government Balance	Deterioration	15% *Current account: –3.7* *Government balance: –1.1*	15% *Current account: –8.2* *Government balance: –1.1*	30%
	Improvement	27% *Current account: –1.9* *Government balance: –4.4*	43% *Current account: –6.6* *Government balance: –5.7*	70%
		42%	58%	100%

Panel B. Distribution of Programs According to the Direction of Envisaged Changes in Government Revenue and Spending[2]

		Expenditure		
		Decrease	Increase	
Revenue	Decrease	30% *Revenue: 28.6* *Spending: 32.0*	14% *Revenue: 26.2* *Spending: 26.2*	44%
	Increase	30% *Revenue: 24.3* *Spending: 30.0*	26% *Revenue: 20.0* *Spending: 24.5*	56%
		60%	40%	100%

Source: MONA database.
[1]Changes are between periods *T*–1 and *T*+1.
[2]Initial levels refer to period *T*–1.

and also the transition group. However, in the case of SBA/EFFs in nontransition countries, two-thirds of the fiscal adjustment was envisaged to come from the revenue side.

- In transition countries the reduction envisaged in the fiscal deficit was milder than average, but in these cases there was a significant reduction in both revenue and spending ratios, reflecting the fact that reduction in the size of the state was also an important objective.[1]

[1]This was itself a response to high levels of both revenue and spending; see Table 2.1.

The averages described above conceal considerable within-group variation that is potentially important for our analysis. The conventional image of IMF-supported programs is that they attempt to improve both the current account deficit and the fiscal deficit, implying a degree of economic austerity on both counts. However, Table 2.3, which shows the distribution of programs according to the direction of envisaged changes in the fiscal balances (as a share of GDP) from the preprogram year *T*–1 to year *T*+1, suggests a more complex reality.

- The current account balance was projected to improve (the current account deficit to narrow) in about 60 percent of programs, but in the re-

maining 40 percent of cases, the current account deficit was projected to widen. The data also show that the direction of change is highly correlated with the initial imbalance; reductions in the current account are associated with large initial deficits, and vice versa.

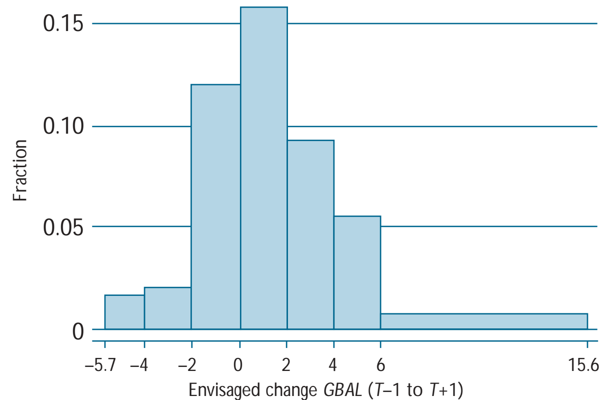
- Overall fiscal balances were envisaged to improve (fiscal deficits to narrow) in 70 percent of cases. In the other 30 percent of cases IMF-supported programs envisaged a widening of the fiscal deficit. In terms of primary balances, the percentage envisaging a widening was even larger, at 35 percent. The conventional view that IMF-supported programs invariably involve fiscal austerity therefore needs some modification. Again, the envisaged direction of change reflects the size of the initial imbalance; the average initial fiscal deficit in the case of programs where the deficit is expected to be reduced is four times larger (as a percent of GDP) than in situations where that deficit is envisaged to widen.

- The composition of the fiscal adjustment in terms of the relative role of revenue increases and spending reductions also varies considerably across countries. Contrary to the perception that IMF-supported programs typically involve a contraction in expenditure as a percentage of GDP, the data show that in 40 percent of cases, total public spending as a percentage of GDP was actually targeted to increase (primary expenditures were projected to increase in 36 percent of the cases). On the revenue side, while about half of the programs envisaged an increase in revenue as a percent of GDP, the other half envisaged a decline.

- The direction of change in expenditures and revenues responds to the initial level of revenue and spending. Programs typically project reductions in spending as a percentage of GDP when initial spending levels are relatively high, and vice versa. Similarly, increases in revenue as a percentage of GDP are envisaged when initial revenue levels are low, and vice versa.[2]

The extent of variation in the targeted fiscal adjustment can be seen from Figure 2.1, which shows the distribution of programs according to the magnitude of the adjustment between *T*–1 and *T*+1. In about two-thirds of programs, fiscal balances are targeted to deteriorate or to improve by less than 2 percent of GDP over a two-year period. The targeted adjustment exceeds 4 percent of GDP in 20 percent of the programs.

Figure 2.1. Distribution of Programs According to the Magnitude of the Envisaged Change in the Overall Fiscal Balance (*T*–1 to *T*+1)
(In percent of GDP)



Source: MONA database.

The phasing of the targeted fiscal adjustment

The phasing of the envisaged fiscal adjustment during the first two years of the program is also of interest. Table 2.4 shows envisaged changes in fiscal balances, and its components, between the preprogram period *T*–1 and each of the two subsequent years *T* and *T*+1.

- On average, programs target a fiscal adjustment of about 1 percentage point of GDP across all types of arrangements during the first year of the program. This figure seems quite stable across different subgroups. Except for the transition economies, this represents between one-half and two-thirds of the total fiscal adjustment over a two-year period.

- In the transition countries all the fiscal adjustment took place in the first year of the program. However, this was also the result of having a lower envisaged fiscal adjustment over a two-year period.

- The different role that expenditure and revenue adjustments are expected to have as the program is implemented is particularly marked in the case of SBA/EFF arrangements in nontransition countries. In fact, spending was not envisaged to decline but rather to increase in the first year,

[2]The fact that program targets respond to initial levels of revenues and expenditures has been documented by IMF staff, including in Abed and others (1998) and Schadler and others (1995a and 1995b).

Table 2.4. The Pace of the Envisaged Fiscal Adjustment
(In percent of GDP)

	Changes in Fiscal Balances from *T*–1 to:	
	T	*T*+1
All arrangements		
Change in fiscal balance	1.1	1.7
Change in revenue	0.8	0.4
Change in expenditure	–0.3	–1.2
ESAF/PRGF arrangements		
Change in fiscal balance	1.1	1.6
Change in revenue	0.6	0.4
Change in expenditure	–0.5	–1.2
SBAs and EFFs (nontransition countries)		
Change in fiscal balance	1.0	2.0
Change in revenue	1.6	1.3
Change in expenditure	0.6	–0.7
SBAs and EFFs (transition countries)		
Change in fiscal balance	1.0	1.1
Change in revenue	–1.2	–1.7
Change in expenditure	–2.2	–2.8

Source: MONA database.

being offset by robust revenue performance to bring about a reduction in the fiscal deficit. The expected relative contributions of revenue and spending are sharply reversed during the second year of the program when spending reductions become more important.

To summarize, the broad conclusion emerging from our examination of 133 arrangements is that IMF-supported programs show a wide variation in the extent of fiscal adjustment, with 30 percent of the arrangements actually projecting a widening of fiscal deficits. In the nontransition cases, programs also incorporate a measure of gradualism in fiscal targets with one-half to two-thirds of the total fiscal adjustment in a two-year horizon being projected to take place during the first year. Furthermore, programs rely relatively less on expenditure adjustment than revenue adjustments (both as a share of GDP) during the first year of the program.

Factors Determining the Scale and Nature of Fiscal Adjustment

The considerable variation in the size of the fiscal adjustment across programs suggests that the adjustment built into an IMF-supported program is not based on some simple mechanical rule of a one-size-fits-all variety. However, the fact that there is variation across countries does not establish that the variation reflects careful calibration of the scale of the fiscal adjustment to the circumstances of each country.

Cross-country regression analysis provides some indication of possible links between the projected fiscal adjustment built into the programs and some of the macroeconomic variables which could be viewed as determinants. Using the fiscal adjustment envisaged over a two-year period, that is, from *T*–1 to *T*+1 as the dependent variable, we have experimented with a number of potential explanatory variables, including the initial size of the fiscal deficit at *T*–1, the size of the current account adjustment envisaged over the period, the initial size of the current account balance, the projected growth rate, the initial level of government spending as a percentage of GDP, and the envisaged change in reserves and inflation.

A complete presentation of the regression results can be found in Appendix 1, Table A1.1. The following estimated equation gave the best fit and all the variables included, except for envisaged growth rate at *T*+1, have coefficients that are statistically significant at the conventional levels.[3] All macroeconomic balances are expressed as a percentage of GDP.

[3] $\Delta GBAL$ = envisaged fiscal adjustment from *T*–1 to *T*+1. $GBAL_{T-1}$ = government balance at *T*–1. CAB_{T-1} = current account balance at *T*–1. ΔCAB = projected change in current account deficit from *T*–1 to *T*+1. EXP_{T-1} = government spending at *T*–1. TR = dummy for transition countries. $Growth_{T+1}$ = envisaged growth rate at *T*+1. The equation was estimated by OLS with heteroskedasticity-consistent standard errors.

$$\Delta GBAL = -1.22 - 0.46 GBAL_{T-1} + 0.11 CAB_{T-1}$$
(–1.51) (–8.52) (2.05)

$$+ 0.18 \Delta CAB + 0.07 EXP_{T-1} - 2.1 TR$$
(4.28) (2.53) (–3.26)

$$- 0.25 TR^{*} GBAL_{T-1} + 0.05 Growth_{T+1}$$
(–1.93) (0.45)

R-squared = 0.61
N = 143

The regression explains 61 percent of the variation in the envisaged fiscal adjustment and the results are quite similar when using the primary fiscal balances as the dependent variable. The main conclusions are the following:

- The most robust finding was a negative association between the size of the programmed fiscal adjustment and the initial (preprogram) level of the fiscal balance. This can be called a tendency toward "fiscal correction": the higher the level of the initial fiscal deficit (or the smaller the fiscal surplus) the stronger is the targeted improvement in the fiscal balance.
- There is a significant positive association between the targeted fiscal adjustment and the envisaged improvement in the current account. In other words, projected improvements in the current account deficit are associated with projected improvements in the fiscal deficit. One can call this a measure of "burden sharing" by the public sector, since the envisaged adjustment in the current account deficit must be shared between the public and private sectors.
- The estimated average "fiscal correction" coefficient for all nontransition arrangements was about –0.5 between (T–1) and (T+1). This implies a reduction of initial fiscal deficits by 50 percent. In the case of transition countries, the fiscal correction coefficient was over –0.70.
- The "burden-sharing coefficients" for all arrangements was about 0.2 for T+1. This means, for example, that projected reductions in the current account deficit of 1 percent of GDP (over the two-year period) are associated with targeted reductions in the fiscal deficit equal to 0.2 percent of GDP over the same period. In other words, only one-fifth of the targeted external adjustment is borne by the public sector. Conversely, if the program envisages a widening in the current account deficit by 1 percent of GDP, it permits a relaxation of the fiscal deficit target by 0.2 percent of GDP.[4]
- The proposed fiscal adjustment is significantly positively associated with the level of expenditures in relation to GDP in the precrisis year. In other words, where expenditure ratios are higher, the fiscal adjustment proposed is larger, a relationship which can be justified because it can be argued that high levels of expenditure also have a crowding-out effect independent of the level of the fiscal deficit.
- Growth assumptions in T+1 were not found to have a significant effect on the targeted fiscal adjustment.
- We found no major difference in these findings across different types of arrangements except for transition countries where, as noted above, the fiscal correction coefficient was larger.

An important limitation of the regression analysis is our inability to test the importance of preprogram public debt ratios as determinants of fiscal adjustment, owing to the absence of comparable data on public debt ratios in the MONA database. This is an important lacuna in the database, which should be corrected for the future.

[4]This symmetry in interpreting the coefficient was tested independently by introducing a dummy variable distinguishing between situations when the current account adjustment was positive or negative.

Chapter 3 Fiscal Adjustment as Presented in Program Documents and the Internal Review Process

In this chapter, we turn to program documents submitted to the Board to consider what light they shed on the rationale for and magnitude of the fiscal adjustment proposed in programs. We then examine the internal review process prior to Board approval as well as during program implementation.

Fiscal Adjustment in Program Request Documents

Why would this be an important question? It could be argued that as long as the substantive aspects of program design have been vetted internally and with the country authorities, there is no reason to worry too much about presentational issues in program documentation. In the view of this evaluation, however, such presentation is indeed important. First, it allows the outside world—stakeholders in the country and the IMF constituency at large, as well as critics—to understand better the rationale for the program, assumptions being made, and why certain measures are taken and not others. It will help the institution convey the message that it has a coherent view that is country and program specific. Second, it may in itself improve program design: the more explicitly the program is explained the more careful the process of internal vetting will have to be.

To address these issues, we examined program request documents of 13 SBA and EFF arrangements.[1]

- Do documents clearly discuss the motivation for the program? Do they explain the nature of the balance of payments problem the program is trying to correct?
- Does the documentation discuss the program-specific mechanism through which the envisaged fiscal deficit adjustment will assist in solving or preventing the external imbalances described above?
- Do documents explain the rationale for the magnitude and pace of the fiscal adjustment and how it is linked to other aspects of the program such as projections for the recovery of growth and private sector activity?
- If there are other factors (besides balance of payments considerations) affecting the need, size, and pace of the envisaged fiscal adjustment, do documents discuss these factors with a reasonable degree of analysis and detail?
- Finally, the IMF has often been criticized for paying much more attention to the magnitude of the needed fiscal adjustment than to the composition of the adjustment. We ask: do documents discuss specific reasons for the distribution of the fiscal adjustment between revenue and expenditure measures?[2]

The methodology used to answer these questions is necessarily subjective: we have reviewed program documents presented to the Board and scored them in each category as "highly satisfactory," "satisfactory," "marginally satisfactory," and "unsatisfactory." Box 3.1 summarizes the criteria used in the scoring (they are further elaborated in Appendix 2, where the code book used in the scoring is presented). This method of evaluating programs depends on subjective judgments of the evaluation team and this is an unavoidable limitation, which must be kept in mind. Nevertheless, we believe the exercise provides useful information to identify some basic patterns. It must also be emphasized that in this analysis we have not focused on

[1]This analysis excludes two ESAF/PRGF arrangements as the fiscal adjustment in these programs was part of a longer-term development strategy rather than a response to shorter-term balance of payments problems as is more typical in SBA/EFF arrangements. The majority of the 13 SBA/EFF programs had a programmed fiscal adjustment above the average of the large sample.

[2]Other qualitative issues of the fiscal adjustment, such as its impact on the equity, efficiency, and sustainability of public finance, are dealt with in other parts of the report.

Box 3.1. How Well Do Documents Explain the Rationale for Fiscal Adjustment?

The five questions raised in considering how well program documents explain the rationale of fiscal adjustment and the criteria used for rating program documents within these dimensions are as follows:

1. Do documents explain the source of the balance of payments problem the program is trying to correct?

We expected documents to provide a coherent and detailed explanation of the sources of the existing or impending external imbalance that the program aimed to correct or prevent. Balance of payments deficits stem from an excess of domestic absorption over income. However, balance of payments problems arise only when such a gap cannot be financed. The table below classifies the various ways the balance of payments problem can be created. When it is due to increases in excess absorption (rooted in the public or private sector), it gives rise to current account problems while financing problems are reflected in the capital account.

2. The link between the balance of payments problem to be corrected and the need for fiscal adjustment.

Does the documentation clearly explain the program-specific mechanism through which the envisaged fiscal deficit adjustment will assist solving/preventing the external imbalances stemming from the specific sources described above? This explanation becomes particularly important if the origin of the balance of payments problem was not directly fiscal (i.e., it emerged from somewhere other than the public sector quadrants of the table below).

3. The magnitude of the fiscal adjustment.

We expected documents to link the magnitude of the fiscal adjustment to the magnitude of the external adjustment, thus explicitly indicating the portion of the total external adjustment borne by the public sector. What is the basis for the "burden sharing" between the adjustment in the private and public sector? This is particularly important in cases where the program envisages an increase in net external financing (deterioration of the current account balance) while envisaging a reduction in the fiscal deficit. What assumptions are being made about the factors that may induce a reduction in the savings-investment balance of the private sector?

4. Other factors influencing the magnitude of the fiscal adjustment.

Programs may incorporate other factors (outside of the ongoing/impending external imbalance triggering the program) in deciding the envisaged fiscal deficit adjustment. It may include reducing inflation if the original deficit was monetized, reducing the crowding out of the private sector, etc. In these cases, do documents clearly discuss how these factors influence the fiscal adjustment?

5. The composition of the fiscal adjustment.

Do documents thoroughly explain the reasons for the envisaged balance between revenue and public spending changes? Are they related to initial levels and efficiency or equity considerations? Are they influenced by the need for speed and expediency?

Origin of External Financing Gap

	Current Account		Capital Account	
	Domestic overheating	External shocks	Rollover problems	Shocks to the supply of net financing
Public sector	Example: increased fiscal deficits due to electoral cycle.	Example: terms of trade shock adversely affecting government revenue.	Rollover and default risk or crisis due to perceived unsustainability of debt (both public and total external) because of country-specific developments.	International contagion affecting private and official flows.
Private sector	Example: private sector consumption and spending bubble financed by borrowing.	Example: terms of trade shock adversely affecting private income or cost of imports.		

Note: Program documents were evaluated on whether they analyzed the sources of the balance of payments problem in these terms.

judging the appropriateness or otherwise of the fiscal adjustment envisaged, but rather on whether the rationale for the adjustment proposed was clearly explained. To assess the appropriateness of the fiscal adjustment proposed in each case would have required an in-depth case study of each country, including an analysis of the combined effect of fiscal and other policies that is beyond the scope of this study.

Table 3.1. Degree to Which Program Documents Explain the Rationale, Magnitude, and Composition of the Envisaged Fiscal Adjustment[1]

	Rating			
	H	S	M	U
1. Does the document clearly discuss the motivation for the program (e.g., the sources of economic disequilibria (balance of payments or other) that the program is expected to address)?		Romania Algeria Ecuador Philippines	Pakistan Peru Ukraine Costa Rica Uruguay	Bulgaria Egypt Jordan Venezuela
2. Do documents explain the country-specific mechanism by which fiscal adjustment will contribute to address actual or potential balance of payments problems?	Ecuador	Bulgaria Philippines Romania Venezuela	Uruguay Algeria Egypt Jordan	Pakistan Peru Ukraine Costa Rica
3. Do documents discuss how the pace and magnitude of fiscal adjustment is being set in order to address the actual or potential balance of payments problems?		Venezuela	Ukraine Costa Rica Romania	Algeria Bulgaria Ecuador Egypt Jordan Pakistan Peru Philippines Uruguay
4. If there are other major factors affecting the envisaged fiscal deficit adjustment (other than balance of payments considerations), do documents explain clearly how they influence the magnitude of that adjustment?	Romania Venezuela Philippines	Bulgaria Egypt Uruguay Algeria	Ecuador Jordan Peru Ukraine	Costa Rica Pakistan
5. Do documents explain the rationale for the composition of the fiscal deficit adjustment, for example, between revenue increases and spending reductions?	Bulgaria Romania Ukraine Algeria	Egypt Jordan Uruguay Philippines Venezuela	Costa Rica Pakistan Peru Ecuador	

Source: Program documents.
[1]Countries listed refer to the program in question. The ratings are: H = highly satisfactory, S = satisfactory, M = marginally satisfactory, and U = unsatisfactory (Appendix 2 contains the code book used for these ratings).

The main results from the analysis are presented in Table 3.1. They can be summarized as follows:

(1) About two-thirds of the programs have been classified as "unsatisfactory" or "marginally satisfactory" in terms of the discussion regarding the justification for IMF involvement or the need for a program in the first place. Most program request documents do not provide either an explanation or a sufficient discussion of the balance of payments problem (current or potential) calling for an IMF-supported program. Where external imbalances do not seem to be the main reason justifying an arrangement, documents often do not present a convincing case of why the arrangement is necessary/recommended on other grounds.[3] Most contain a background section with recent economic developments, but the scope and degree of analysis varies greatly across programs. Some Board papers only provide one historical paragraph and take for granted the reasons why the program is needed.

(2) Even if it is accepted that in most cases the motivation for the program was to address external imbalances, a majority of the programs reviewed did not explain the links between the targeted fiscal adjustment and the envisaged improvement in the external situation. In only 40 percent of cases was there an explicit discussion of the program-specific mechanism through which fiscal adjustment could help improve the external imbalance.

(3) Even when the Board papers identify the link between fiscal adjustment and external adjustment, the documentation does not discuss how the specific pace and magnitude of the fiscal adjustment is being set to attain the new balance of payments situation envisaged in the program. This is the area with the worst scores: 9 out of the 13 cases were rated unsatisfactory. Only in one case was it satisfactory.

[3]Program objectives are often described in very general terms such as "restoring and sustaining a high rate of economic growth," "alleviating inflationary pressures," or "re-establishing balance of payments viability over the medium term." What is usually missing, however, is a discussion of why these issues require the involvement of the IMF and the disbursement of resources in each particular case.

(4) When there were other factors (other than balance of payments considerations) affecting the envisaged fiscal adjustment, only half of the programs clearly explained how these factors influenced that adjustment. Furthermore, only in three cases (those judged as highly satisfactory in Table 3.1) was there an explanation of how these factors influenced the magnitude of the fiscal adjustment.

(5) Most programs provide a good explanation for the composition of the envisaged fiscal adjustment between revenue increases and spending reductions. In some cases, the analysis is very good indeed, providing not only a sense of why fiscal adjustment should be distributed in the proposed way, but also analyzing intra-revenue and intra-spending changes that can help improve the structure of public finance. However, in about one-third of the cases, programs provide only a long list of measures on both the revenue and spending side without a sufficient sense of rationale or priority.

Fiscal adjustment and debt sustainability

An area that is conceptually important, but that received less attention than it deserved in the sample of program documents examined earlier, was the linkage between the fiscal deficit and debt sustainability. Table 3.2 provides data on the average ratio of public debt to GDP for 12 of the 13 countries for which IMF-supported program documents were examined in the last section.[4] During the preprogram years, debt ratios were relatively low for Ukraine, Romania, and Uruguay but were relatively high in the other countries. Nevertheless, only five programs included a discussion linking the dynamics of public debt to the targeted fiscal adjustment, and even then the linkage to the magnitude of fiscal adjustment needed was weak. This is clearly an area where practice needs to be greatly strengthened and analyses made more explicit. This lacuna has been recognized and more recent practice seems much better.[5]

Table 3.2 shows, for purposes of comparison, what happened to the debt ratios. The picture is mixed. Although debt ratios were reduced in half of the programs, in one case, Ecuador, it was largely due to debt-reduction initiatives during the program years.

These short-term changes are not necessarily a good indication of sustainability and robustness of fiscal systems over the medium term because they are influenced by short-term fluctuations in exchange rates, interest rate premia on public debt, and the like. Even if debt ratios rise in the short run, long-term sustainability may be improving if the higher debt levels reflect the short-term costs of reforms that yield benefits over the medium term. Conversely, short-term improvements may not be sustainable if they are achieved through short-term fiscal deficit reduction measures that quickly erode over time (e.g., increasing already high taxes over a low base therefore inducing further tax evasion, or unsustainable cuts in public sector wages without civil service reforms). These problems can be handled satisfactorily only through systematic debt-sustainability analysis which takes account of the different factors that affect debt profiles over time, in particular structural reforms in fiscal systems, an area to which we return later on.

Conclusions

Our evaluation suggests that although there is considerable variation across countries in the direction and size of the fiscal adjustment proposed in IMF-supported programs, there is not enough clarity and transparency in the way fiscal targets are set. Program documents should indicate clearly the extent to which fiscal deficit adjustments proposed are being driven by consideration of burden sharing in reducing aggregate demand, debt-sustainability considerations, or crowding-out concerns.

The Internal Review Process Prior to Board Approval and During Program Implementation

In addition to examining program documents we also attempted to evaluate the extent to which the internal review process for the 15 sample programs focused on key areas of public finance reform. We selected three broad areas: (1) macro and fiscal targets; (2) specific program design issues in the fiscal area; and (3) equity and social spending issues.

Comments from PDR and FAD were reviewed at three stages: (1) the briefs for negotiation; (2) Letter of Intent (LOI) or Memorandum of Economic and Financial Policies (MEFP) and staff reports for the initial program request; and (3) subsequent reviews during program implementation. In the following, a reference to "brief" indicates comments on the brief for the negotiating mission; "LOI/MEFP and Staff Report" deals with comments on the LOI/MEFP or Staff Report for the initial program request and "Reviews" covers remarks on all documents (briefs, LOI/MEFP, and staff reports) related to reviews during program implementation.

[4]Egypt is not included because of lack of data on domestic public debt during the relevant period.

[5]Staff reports are increasingly assessing debt sustainability following the framework outlined in IMF (2002e).

Table 3.2. Average Public Debt Prior to and Following the Initial Program Year

Country (Program)	Average *T*–3 to *T*–1	Average *T*+1 to *T*+3	Difference
Algeria (SBA 1994)	88.0[1]	81.8	–6.2
Bulgaria (EFF 1998)	112.1[2]	78.3	–33.8
Costa Rica (SBA 1995)	49.1	50.5	1.4
Ecuador (SBA 2000)	76.7	64.9[3]	–11.8
Jordan (EFF 1999)	91.7	66.7	–25.0
Pakistan (SBA 2000)	93.1	95.9[4]	2.8
Peru (EFF 1996)	50.5	41.3	–9.2
Philippines (SBA 1998)	108.5	128.0	19.5
Romania (SBA 1999)	27.9[2]	28.7[3]	0.8
Ukraine (EFF 1998)	30.1	44.7	14.6
Uruguay (SBA 2000)	28.1	63.1[3]	35.0
Venezuela (SBA 1996)	63.7	30.6	–33.1

Source: IMF staff estimates.
[1]No data available for *T*–2 and *T*–3.
[2]No data available for *T*–3.
[3]No data available for *T*+3.
[4]No data available for *T*+2 and *T*+3.

Overview of findings

Table 3.3 summarizes the findings by identifying the countries for which each issue listed was raised. Comments on the brief for the initial program request were largely concentrated on the macroeconomic framework and fiscal targets and covered 11 out of 15 programs. Comments on the brief regarding program design were limited to 6 programs and those dealing with equity and social spending extended to 7 programs. Taking all comments at every stage combined, reviewers commented on the macroeconomic framework and fiscal targets in all 15 programs while the other two topics were commented on in only 11 of the 15 programs.[6]

The area with the most coverage (over the program life cycle) was the discussion of risks to achieving the fiscal targets. This surfaced in only 4 program cases (Costa Rica, Egypt, Pakistan, and Romania), at the time of reviewing the initial brief, but over the life of the program coverage expanded to 13 of the 15 countries, and comments were often extensive. Other areas that were well covered (12 programs) include concerns on fiscal sustainability and suggestions to discuss the factors behind the magnitude and pace of the fiscal adjustment. While initial focus on fiscal sustainability was limited to only 4 programs, reviewers pressed on the rationale for the magnitude and pace of adjustment in 9 of the 15 briefs.

A potentially important feature emerging from the evaluation is the phenomenon of a larger concentration of comments in the later stage reviews when they can only affect mid-course correction rather than initial program design. For all three topics (macroeconomic and fiscal targets, program design, and equity and social spending) there were more comments during the program review phase than at the initial briefing stage. Yet, it is in the earlier stages that comments could have the most influence on the nature of the adjustment process and the biggest impact on the fiscal adjustment strategy.

This finding is particularly surprising since many of the later comments relate to basic design issues, such as risks to achieving the fiscal target, fiscal sustainability, the size of the deficit, optimism of growth projections, and the importance of embedding the adjustment path in a medium-term outlook. However, for many of the programs, these issues are either not raised at all or only lightly touched upon at the stage of the initial negotiation brief.

[6]For 3 out of the 15 cases (Senegal, Ukraine, and Uruguay) there were only a few comments in the fiscal area and these were concentrated on the review phase. There were comments dealing with other areas than those we focus on here. For example, in the case of the Senegal brief, comments focused on energy subsidies, crop credit, and HIPC-related issues. Moreover, FAD agreed with the 1998 country strategy paper and therefore saw no need for major comments on the negotiation brief. For the Ukraine brief, reviewers emphasized the ramifications of implementing measures by decree, the slow progress with overall reform efforts, and the implications of unbudgeted payments to the coal sector. Furthermore, the 1998 Ukraine EFF was negotiated over a long period of time with several comments relating to the negotiation and design of the program having been communicated before the negotiating mission brief was prepared. As a result, there was less need to comment on the initial brief. FAD involvement in the Uruguay program became more important at later stages when the economic crisis intensified and a follow-up SBA was requested.

Table 3.3. Selected Topics Commented on During the Review of 15 IMF-Supported Programs

	Brief for Negotiation of the Arrangement	LOI/MEFP and Staff Report for the Program Request	Documents Relating to Reviews of the Arrangement
Macro and fiscal targets			
Were questions raised on realism of GDP growth projections?	Egypt, Romania	Egypt	Bulgaria, Pakistan, Philippines, Romania, Uruguay
Were questions raised on realism of projected private sector investment response?	Costa Rica, Egypt, Jordan	Bulgaria	Algeria, Bulgaria, Egypt
Any suggestions to link growth/private sector recovery and fiscal adjustment?	Bulgaria	Peru, Philippines	Bulgaria, Philippines
Any suggestions to discuss the factors behind the magnitude and pace of the fiscal adjustment?	Bulgaria, Ecuador, Jordan, Pakistan, Peru, Philippines, Romania, Tanzania, Venezuela	Algeria, Bulgaria, Ecuador, Egypt, Peru, Venezuela	Algeria, Bulgaria, Ecuador, Egypt, Jordan, Peru, Philippines, Romania, Senegal, Tanzania, Venezuela
Any discussion of risks to achieving fiscal targets?	Costa Rica, Egypt, Pakistan, Romania	Costa Rica, Ecuador, Jordan, Pakistan, Peru, Philippines, Romania, Tanzania, Venezuela	Algeria, Bulgaria, Ecuador, Jordan, Peru, Philippines, Romania, Tanzania, Uruguay
Any suggestions to disentangle causes of fiscal underperformance?	Jordan	Bulgaria, Philippines, Romania	Bulgaria, Philippines
Any concerns raised on fiscal sustainability?	Bulgaria, Ecuador, Pakistan, Peru	Algeria, Bulgaria, Ecuador, Jordan, Peru, Romania, Tanzania	Algeria, Bulgaria, Ecuador, Egypt, Jordan, Pakistan, Peru, Philippines, Romania, Tanzania, Ukraine, Uruguay
Program design			
Any comments regarding the need to consider past implementation problems in designing the program?	Bulgaria, Jordan, Pakistan, Peru	Jordan, Pakistan, Peru, Tanzania	Egypt, Jordan, Pakistan, Peru, Philippines, Ukraine, Venezuela
Any suggestions for stronger action in areas under the control of the executive that would demonstrate political commitment to the program?	Bulgaria, Ecuador, Jordan, Peru, Tanzania	Bulgaria, Ecuador, Jordan, Pakistan, Philippines, Romania, Tanzania	Bulgaria, Ecuador, Egypt, Jordan, Pakistan, Peru, Philippines, Tanzania
Any comments urging a more realistic time frame for reform?	Bulgaria, Jordan, Pakistan		Bulgaria, Pakistan
Equity and social spending			
Any suggestions to deal with issues that would improve the equity of taxation?	Bulgaria, Egypt, Pakistan	Bulgaria, Egypt, Jordan, Peru, Venezuela	Bulgaria, Ecuador, Philippines, Tanzania
Any suggestions to deal with issues that would improve the equity of spending?	Egypt, Tanzania	Egypt, Jordan	Bulgaria, Ecuador, Tanzania
Did the review process raise the need to analyze trends in social spending?	Algeria, Bulgaria, Jordan	Algeria, Bulgaria, Egypt, Jordan, Pakistan, Tanzania	Algeria, Bulgaria, Ecuador, Egypt, Pakistan, Peru, Tanzania
Any proposals for staff to identify how social spending could be protected?	Algeria, Egypt, Jordan, Venezuela	Ecuador, Jordan, Pakistan, Philippines, Venezuela	Algeria, Ecuador, Egypt, Pakistan, Peru, Philippines, Romania, Venezuela

Source: Program documents.

Half the programs received no substantive comments on strategic elements of design in the initial negotiating brief. For the other half, there was an equal split between those with substantive and detailed comments (Bulgaria, Ecuador, Pakistan, and the Philippines) and those where reviewers touched on design issues more lightly (Egypt, Jordan, Peru, and Venezuela). It may be significant that 3 of the 4 pro-

grams with stronger comments on program design issues were follow-up programs (Bulgaria, Pakistan, and the Philippines). Comments dealing with strategic issues at the stage of the negotiating brief concentrated on the need to explain the factors behind the proposed pace, magnitude, and composition of the fiscal adjustment (9 out of 15 programs). Suggestions for stronger action that would demonstrate political commitment to the program receive the next amount of attention at this early stage but were limited to only one-third of programs. Other issues were raised in less than one-third of the programs.

The range of issues covered under program reviews (as the program was being implemented) became much wider. Concerns on fiscal sustainability were expressed in 12 programs. In at least half the programs, comments were made in areas such as pace, magnitude, and composition of adjustment; risks to achieving fiscal targets; actions that demonstrate political commitment; and considering past implementation problems in the design of reforms. For illustrative purposes, we provide some specific examples of the issues raised in these areas.

Test of political commitment

Many specific comments were raised in this area at the review stage. Common themes are the need to roll back exemptions, limit tax incentives, take concrete action against tax evasion, and enforce collection of tax arrears. However, in only three cases (Bulgaria, Jordan, and Peru) did reviewers press for action in these specific items in commenting on the program negotiation brief.

Some of the specific issues raised during the reviews could have been suggested with greater effectiveness at the outset of the program. These include, for example, the removal of tax preferences from the Foreign Investment Act in Bulgaria; concrete suggestions to reduce tax evasion in Ecuador; elimination in Pakistan of exemptions identified by the Committee on Reform of the Income Tax Ordinance; enforcing collection of the largest tax arrears in Peru; bringing forward the rationalization of fiscal incentives in the Philippines; harmonizing Tanzanian income tax relief and investment incentives between the mainland and Zanzibar; and reaching understandings on enforcing bankruptcy laws on major delinquent tax payers in Ukraine. Had these suggestion surfaced at the outset of the program they might have provided an early test of the commitment of the authorities to reform and contributed to a stronger program design.

Revenue issues

Over the life cycle of the program reviewers often pressed for unbundling the reasons for poor tax administration; suggested addressing basic issues of sequencing between reducing some taxes and enlarging the tax base; and reforming income taxes to improve equity as well as effectiveness. Most common were suggestions for reforms to improve the equity of taxation. Comments in this area extended to nine programs. However, coming as they did during the review phase, these comments were perhaps less useful.

Conclusions

The review process raised many questions in critical areas also identified as weak by this evaluation (Chapter 7). However, reviewers ended up in a reactive mode, commenting more extensively as programs proceeded, rather than strategically at the outset, when they could have had the biggest impact to improve program design. This is puzzling since later comments are rich and focus on a wide array of critical design issues. Areas where comments could ideally come at the negotiation phase include (1) the need to link the fiscal adjustment with the recovery of private sector activity and growth; (2) specific actions that would demonstrate the commitment of the authorities to the program, particularly in removing exemptions and taking more forceful actions to reduce tax evasion; (3) consideration of past implementation problems in the design of the program; and (4) measures to improve equity and protect social spending.

The tendency for comments to come late in the process can be explained by reference to several important factors:

- Reviewing departments seem to provide latitude to mission chiefs in developing the details of the program, particularly when projected outcomes appear ambitious. Subsequent comments reflect the fact that developments are less favorable than originally expected.
- Program briefs basically address the immediate requirements of the economy for the early stage of adjustment, while other complementary policies are needed only for later stages. Reviews would then follow up on this sequencing of measures, with the functional departments possibly taking the lead.
- There are time constraints associated with the preparation of prenegotiation briefs, which make it difficult to achieve more substantive comments at the outset.

These are important considerations but there is need for a concerted effort to ensure more brainstorming and pooling of potential review resources at an earlier stage when it can have a stronger ef-

fect on program design. This could be helped by having the brief articulate more clearly the basis for the fiscal program, including a medium-term fiscal scenario incorporating a basic debt sustainability discussion. This sequence would also be more consistent with the greater emphasis on domestic ownership, since it would involve a fuller exploration of alternative policy options, and potential trade-offs, at the initial design stage. The new guidelines on the review process issued by the First Deputy Managing Director on March 30, 2003 call for early consultations across departments in order to form a common understanding of the issues that are subsequently addressed in the papers. Such earlier consultations would help improve program design.

Chapter 4 Fiscal Performance Compared with Targets

In this chapter we compare actual fiscal performance against targets projected in programs. We first examine the large cross section of programs to assess the frequency and nature of fiscal shortfalls relative to targets, the sources of these shortfalls, and the extent to which programs have been flexible in revising targets as the program unfolds. We then examine in some depth the 15 sample programs to study the composition of the fiscal adjustment and some qualitative dimensions of that adjustment that cannot be detected in the cross-country analysis.

Cross-Country Analysis

Table 4.1 compares actual with envisaged (average) changes in the current account and fiscal balances over a two-year horizon for the sample as a whole and also for the different subgroups. The following broad patterns emerge:

- Whether we look at the overall balance or the primary balance, fiscal balances improved by half the projected amounts for the sample as a whole. Shortfalls relative to projections were about ¾ of 1 percent of GDP in both cases. However, there were important differences across subgroups. Fiscal targets were met in the transition countries but not in the other subgroups. Concessional arrangements and SBA/EFFs in nontransition countries experienced a shortfall in the primary balance equal to half the targets, although these shortfalls were small as a percentage of GDP (0.4 percent and 1 percent, respectively).
- The composition of the fiscal adjustment also shows significant variation across types of programs. In the case of SBA/EFF arrangements in nontransition countries, the shortfall is largely due to the expenditure side: while programs projected on average a reduction in public expenditures to the tune of 0.7 percent of GDP, expenditures actually increased by 0.6 percent of GDP. In the case of SBA/EFF arrangements in transition countries, both revenue and expenditure declined more or less in line with projected values.
- While fiscal balances improved much less than projected, the current account balances on average improved slightly more than projected, though the pattern varied across subgroups. In concessional programs, the actual developments turned out to be equal to projections but in the SBA/EFF arrangements, the current account position adjusted more than projected in both subgroups (nontransition countries and transition countries). In the nontransition countries the average current account adjustment exceeded projections by more than 1 percent of GDP, a statistically significant change.

As noted in other comparisons, the group averages mask considerable variation across countries and this is shown in Table 4.2, which presents the distribution of programs according to their fiscal performance relative to program targets by the second year of the program.

- About 58 percent of programs had a shortfall with respect to targets in the overall fiscal balance, and the percentage in the case of primary balances was 66 percent. The mean fiscal shortfall in this group was 2.8 percent of GDP for the overall balance and 2.2 percent of GDP for the primary balance.
- The incidence was largest in SBA/EFF arrangements in nontransition countries, with about three-fourths of these programs having fiscal shortfalls. It is followed by ESAF/PRGF programs, where about half the arrangements had shortfalls. The mean shortfalls for both these groups are similar, about 3 percent of GDP.
- In contrast, SBA/EFF arrangements in the transition countries had the lowest incidence of shortfalls in overall balances (40 percent). However, the picture changes when the incidence of shortfalls refers to primary balances (60 percent). It is clear that in these arrangements,

Table 4.1. Changes in External and Fiscal Balances from (*T*–1) to (*T*+1)[1]
(In percent of GDP)

	All Arrangements	ESAF/PRGF	SBA/EFF: Transition countries	SBA/EFF: Nontransition countries
Envisaged				
Current account	–0.3	0.1	–2.0	–0.2
Government balance	1.7	1.6	1.1	2.0
Primary balance[2]	1.4	1.0	0.4	2.0
Government revenues	0.4	0.4	–1.7	1.3
Government expenditures	–1.2	–1.2	–2.8	–0.7
Actual				
Current account	0.3	0.1	–1.3	1.1
Government balance	0.8	1.0	1.8	0.2
Primary balance[2]	0.7	0.6	0.3	1.0
Government revenues	0.2	0.1	–1.4	0.9
Government expenditures	–0.7	–1.0	–3.2	0.6
Count	133	60	21	52

Sources: MONA and WEO databases.
[1]Figures subject to rounding errors.
[2]Based on a sample of 115 arrangements.

Table 4.2. Differences Between Actual and Projected Changes in Fiscal Balances[1]
(From T–1 to T+1)

Distribution of Programs (In percent)

Differences (In percent of GDP)	All arrangements	ESAF/PRGF	SBA/EFF: Transition countries	SBA/EFF: Nontransition countries
Positive differences (“overperformance”)				
Larger than 4	4.5	3.3	13.6	2.0
Between 3 and 4	3.7	5.0	4.5	2.0
Between 2 and 3	5.2	6.6	0.0	5.9
Between 1 and 2	10.5	13.3	18.2	4.0
Between 0 and 1	18.2	20.2	22.8	13.8
Subtotal	42.1 (33.8)	48.4 (43.1)	59.1 (41.2)	27.7 (22.4)
Mean	1.9 (2.1)	1.9 (2.2)	2.2 (1.8)	1.7 (2.1)
Negative differences (“underperformance”)				
Between 0 and –1	12.8	11.7	22.8	9.8
Between –1 and –2	18.1	13.3	9.1	27.7
Between –2 and –3	9.0	10.0	4.5	9.8
Between –3 and –4	5.2	3.3	4.5	7.8
Smaller than –4	12.8	13.3	0.0	17.1
Subtotal	57.9 (66.2)	51.6 (56.9)	40.9 (58.8)	72.2 (77.6)
Mean	–2.8 (–2.2)	–2.9 (–3.0)	–1.4 (–1.5)	–3.1 (–2.5)
Total	100.0	100.0	100.0	100.0
Count	133	60	21	52
Overall mean	–0.8* (–0.6)*	–0.6 (–0.4)	0.7 (–0.2)	–1.8* (–0.9)*
Std	3.5	3.4	2.6	3.7

Sources: MONA and WEO databases.
*This difference between actual and envisaged adjustment is statistically significant at the 99 percent or better confidence level.
[1]Values in parentheses show the results when overperformance and underperformance are defined in terms of primary balances.

Table 4.3. The Dynamics of Fiscal Adjustment

(In percent of GDP)

		(T–1) to T			(T–1) to (T+1)		
	N	Envisaged	Actual	Difference[1]	Envisaged	Actual	Difference[1]
Changes in fiscal balances							
All arrangements	133	1.0	0.8	–0.2	1.7	0.8	–0.9[2]
SBA/EFF							
Transition countries	21	1.0	1.3	0.3	1.1	1.8	0.7
Nontransition countries	52	1.0	0.5	–0.5	2.0	0.2	–1.8[2]
ESAF/PRGF	60	1.0	0.9	–0.1	1.6	1.0	–0.6
Changes in revenues							
All arrangements	133	0.8	0.4	–0.4	0.4	0.2	–0.2
SBA/EFF							
Transition countries	21	–1.2	–1.2	0.0	–1.7	–1.4	0.3
Nontransition countries	52	1.7	1.1	–0.6	1.3	0.9	–0.4
ESAF/PRGF	60	0.6	0.3	–0.3	0.4	0.1	–0.3
Changes in expenditure							
All arrangements	133	–0.3	–0.4	0.1	–1.2	–0.7	–0.5
SBA/EFF							
Transition countries	21	–2.2	–2.4	0.2	–2.8	–3.2	0.4
Nontransition countries	52	0.6	0.6	0.0	–0.7	0.6	–1.3[2]
ESAF/PRGF	60	–0.4	–0.6	0.2	–1.2	–1.0	–0.2

Sources: MONA and WEO databases.

[1]Difference refers to the actual minus envisaged magnitudes. Hence, negative values show underperformance.

[2]Difference statistically significant at the 95 percent confidence level. The other differences in means are not statistically significant.

programs have frequently underestimated favorable developments in interest payments providing relief to the budget.

The pace of the adjustment

The pace of the fiscal adjustment during the first two years of the program provides some interesting insights. Table 4.3 compares projected and actual changes in fiscal balances between the preprogram year *T*–1 and the program years *T* and *T*+1.

- Almost all fiscal adjustment on average takes place during the first year of the program. Except in the transition countries, programs were unable to achieve further fiscal gains in the second year of the program in spite of more ambitious fiscal targets.
- In SBA/EFF arrangements in nontransition countries, revenue ratios did not increase beyond the gain of 1 percentage point of GDP achieved during the first year of the program and expenditure ratios could not be reduced.
- Concessional programs exhibited similar features, except that these programs were able to reduce expenditure ratios by the second year of the program. It is possible that this is because financing for these countries was more of a binding constraint than for the other cases.

Composition of the adjustment in programs with fiscal shortfalls

We now turn to examine the anatomy of programs with fiscal shortfalls, namely whether fiscal shortfalls are primarily due to revenues (as a share of GDP) being below target or expenditures (as a share of GDP) above target. The relevant data are presented in Table 4.4.

Except for programs in transition countries, a much larger proportion of programs reflects situations where excess expenditure as a share of GDP (relative to targets) is the dominant source of the fiscal shortfall. This is particularly important in the case of ESAF/PRGF arrangements.

It is relevant to ask whether the shortfalls are the result of very ambitious fiscal targets (on either the revenue or expenditure side) or the result of moderate targets combined with very little progress in the actual adjustment. Table 4.5 presents a comparison of fiscal targets and actual achievements for the group of programs showing a shortfall.

- The fiscal shortfall is largest for the group of nontransition SBA/EFF arrangements, where the fiscal deficit, far from showing an improvement by *T*+1, actually shows a deterioration. However, the volume of adjustment proposed in this group was not larger than for others. In fact, it is the subgroup of transition economies that

Table 4.4. Percentage Distribution of Programs with Fiscal Shortfalls
(Shortfalls expressed as a share of GDP)

	All Arrangements	ESAF/PRGF	SBA/EFF Transition countries	SBA/EFF Nontransition countries	Large Episodes of Envisaged Fiscal Adjustment
Programs where at least half the fiscal shortfall is due to:[1]					
Expenditure shortfalls	72.0	84.0	50.0	67.0	29.0
Revenue shortfalls	28.0	16.0	50.0	33.0	71.0
Total	100.0	100.0	100.0	100.0	100.0

Sources: MONA and WEO databases.
[1]The shortfall is the difference between actual and projected values. For example, if the fiscal deficit is 3 percentage points of GDP higher than programmed, and spending is 1 percentage point of GDP higher than envisaged while revenue is 2 percentage points of GDP lower than projected, spending accounts for one-third of the fiscal shortfall and revenue for two-thirds.

Table 4.5. The Composition of Fiscal Adjustment in Programs with Fiscal Underperformance
(In percent of GDP; values in parentheses refer to primary balances or primary expenditures[1])

	N	(*T*–1) to (*T*+1) Envisaged	Actual	Difference[2]
Changes in fiscal balances				
All arrangements	77	2.3	–0.5	–2.8
SBA/EFF				
Transition countries	8	4.3	2.9	–1.4
Nontransition countries	38	2.0	–1.0 (0.0)	–3.0 (–2.0)
ESAF/PRGF	31	2.1	–0.8	–2.9
Changes in revenues				
All arrangements	77	0.5	–0.1	–0.6
SBA/EFF				
Transition countries	8	–2.7	–2.4	–0.3
Nontransition countries	38	1.4	0.5	–0.9
ESAF/PRGF	31	0.3	–0.3	–0.6
Changes in expenditures				
All arrangements	77	–1.8	0.4	–2.2
SBA/EFF				
Transition countries	8	–7.1	–5.3	–1.8
Nontransition countries	38	–0.6	1.6 (0.6)	–2.2 (–1.2)
ESAF/PRGF	31	–1.8	0.5	–2.3

Sources: MONA and WEO databases.
[1]Values for primary balances or primary expenditures are presented only if they significantly differ from overall fiscal balances or total expenditures.
[2]Difference refers to the actual minus envisaged magnitudes. Hence, negative values show underperformance.

shows the highest proposed improvement and this group also had the best compliance record.

- In the case of SBA/EFF in nontransition countries, about two-thirds of the adjustment was expected to come from the revenue side and one-third from expenditure. In fact, revenues increased much less than expected while expenditures increased in relation to GDP, instead of declining as programmed.

- In the ESAF/PRGF programs, both revenues and expenditures moved in the opposite direction compared with projections. Revenues declined originally instead of increasing, and expenditures increased instead of declining as

Table 4.6. Changes in Government Balances in Large Episodes of Envisaged Adjustment from (*T*–1) to (*T*+1)
(In percent of GDP)

		Government Balance			Revenues			Expenditures		
	N	Envisaged	Actual	Shortfall	Envisaged	Actual	Shortfall	Envisaged	Actual	Shortfall
Total	39	5.7	3.6	–2.1[1]	1.7	–0.5	–2.2[1]	–4.0	–4.1	0.1
SBA/EFF										
Transition countries	6	6.5	5.2	–1.3	–2.1	–5.7	–3.6	–8.6	–10.8	2.2
Nontransition countries	17	5.4	3.4	–2.0	3.0	1.5	–1.5	–2.4	–2.0	–0.4
ESAF/PRGF	16	5.7	3.1	–2.6[1]	1.7	–0.8	–2.5[1]	–3.9	–3.9	0.0

Sources: MONA and WEO databases.
[1]Difference statistically significant at the 95 percent confidence level.

programmed and the latter effect explains most of the shortfalls.

- In the transition countries, the fiscal shortfall is due to expenditure shortfalls, in spite of programs achieving expenditure reductions equivalent to 5.3 percent of GDP. One may argue that this shortfall is to be expected given the significant expenditure cuts being programmed, equal to 7 percent of GDP.

Large episodes of envisaged adjustment

The case of large envisaged fiscal adjustment (defined as larger than 3 percent of GDP between *T*–1 and *T*+1) is of special interest because the results described above are reversed; revenue shortfalls account for most of the cases.[1] The results are shown in Table 4.6.

- The average targeted fiscal adjustment in this subgroup is 5.7 percent of GDP (Table 4.6) compared with only 1.7 percent of GDP for the 133 arrangements taken together (Table 4.3). The initial fiscal deficit in the preprogram year in this group is also higher, at 7.8 percent of GDP compared with 4.1 percent for the overall sample.
- Programs achieve half of the envisaged adjustment; substantial adjustment was undertaken notwithstanding the shortfalls. Much of this adjustment was on the expenditure side. In fact, expenditure reductions for this group were significantly stronger than in milder cases of fiscal adjustment.
- Revenue shortfalls remain significant in these cases, in spite of the higher requirements for revenue increases stemming from the need to reduce a more severe initial fiscal imbalance.

A possible explanation for the heavier reliance on expenditure cuts in programs with very large initial fiscal imbalances is simply that the deficits could not be financed and large expenditure cuts became unavoidable when revenue measures did not yield results quickly enough.

This conclusion has special applicability in SBA/EFF arrangements in nontransition countries. In spite of the relatively higher level of development in these countries, programs were not able to raise more than 1.5 percent of GDP in extra revenues by the second year of the program, irrespective of the severity of the initial fiscal deficit and the size of the targeted fiscal deficit. Expenditures were then cut residually.

Determinants of fiscal shortfalls

Regression analysis is one way of identifying possible determinants of the fiscal shortfalls and our results are presented in detail in Appendix 1, Table A1.2. The most significant variable was the difference between the envisaged and actual rate of growth for *T*+1.[2] Lower-than-envisaged GDP growth

[1]Three extreme cases of major expenditure collapse were excluded from this group: Armenia SBA 1995; Equatorial Guinea ESAF 1993; and Malawi ESAF 1995. In these programs, public expenditures collapsed between 18 and 26 percent of GDP.

[2]It may be surprising that deviations in growth projections from actuals do explain deviations in fiscal adjustment while growth projections did not seem to have influenced a program's fiscal adjustment projections (as found earlier in Chapter 2, "Factors Determing the Scale and Nature of Fiscal Adjustment"). This apparent puzzle is explained by the fact that actual fiscal adjustment was indeed found to be associated with actual growth (Appendix 1, Table A1.1). In fact, the coefficient of the growth variable in the equations for deviations is similar to the one in the equation for actuals. The fact that growth is not significant in the fiscal projection equations does not mean that errors in growth projections do not influence shortfalls in fiscal adjustment. This effect will persist as long as actual fiscal adjustment is influenced by actual growth rates.

was associated with less fiscal adjustment than envisaged; a shortfall in the growth rate with respect to projections equal to 1 percentage point of GDP was associated with a fiscal shortfall compared to programmed levels of 0.3 percentage point of GDP.

It is relevant to ask whether this effect operates via the expenditure or revenue side. To explore these channels, separate regressions were run to explain revenue and expenditure ratio shortfalls. Interestingly, growth optimism proved to be significant in explaining optimism in reducing expenditures as a share of GDP (also with a coefficient equal to 0.3 but negative) but not in explaining optimism in forecasting revenues as a share of GDP. In fact, the elasticities of projected and actual revenues with respect to GDP happen to be similar and close to one, so that shortfalls in GDP growth lead to proportional shortfalls in revenue without much effect on the revenue/GDP ratio.[3] On the other hand, optimism in growth generates optimism in projecting declines in the expenditure/GDP ratios which are not realized in the end because nominal expenditure levels are usually less sensitive to growth.

The low explanatory power of the regressions—they explained only about 22 percent of the variation in the differences between envisaged and actual fiscal adjustment—suggests the omission of important explanatory variables that may also influence revenue ratio shortfalls. This is particularly true in the episodes of large envisaged fiscal adjustment, where programs may have overestimated the speed at which tax policy and tax administration measures could be implemented, or the extent by which these policies could quickly yield revenue increases. The overall role of optimism regarding the progress of structural reforms in the fiscal area is examined later on in Chapter 7.

Flexibility of Fiscal Targets During Program Reviews

In view of the persistent shortfall in fiscal performance compared with targets, it is relevant to consider how programs are adjusted to take account of shortfalls. This section first looks at the magnitude and direction of revisions in fiscal targets in a large sample of programs in the 1993–2001 period. The results are then complemented by a qualitative analysis of the 15 program case studies. We examine three interrelated questions associated with the extent of fiscal flexibility in programs. First, in what direction are fiscal targets being revised? Second, are the revisions linked to changes in other projected variables such as GDP growth? Finally, do program documents provide a good rationale for the revision? To answer these questions we have analyzed how the first and second reviews modified fiscal targets for year $T+1$, namely, the second year of the program.[4]

Patterns in the revision of fiscal targets

Revision in fiscal targets (and indeed in other key macroeconomic targets) is common in the course of program review. Figure 4.1 presents the distribution of differences between the original fiscal target and the revised target in the first review (Panel A) and between the original target and the second review (Panel B).

Panel A shows that at the first review 55 percent of the programs experienced minor revisions (between plus/minus 0.5 percentage point of GDP) and the few cases of large revisions are more or less distributed symmetrically in both directions. These results are not surprising given that the first review is relatively close to the date of the original program request. However, by the time of the second review (Panel B), the center of the distribution shifts to the left suggesting that by the second year many more targets are relaxed (targeting a lower improvement in the fiscal balance) than tightened (targeting a greater improvement in the fiscal balance).

Since fiscal outcomes are affected by growth outcomes, it is relevant to consider whether revisions in fiscal targets in the course of programs reflect revisions in growth projections (Table 4.7). The data suggest that both distributions are indeed related.[5] However, it is interesting to note that growth revisions do not seem to operate in a symmetric way: when growth projections are revised downward, fiscal balance targets are adjusted downward in two-thirds of the cases. By contrast, when growth projections are revised upward, fiscal targets are adjusted upward in only half of the cases.

This asymmetry was also captured when we ran separate regression equations for the cases where

[3]Regressions to explain both envisaged and actual changes in revenue/GDP confirm that forecast and actual revenue-GDP elasticities are similar and close to one.

[4]We chose $T+1$ rather than T for two reasons. First, very few programs have more than one revision during the first year of the program. Hence, looking only at one revision would provide a partial and incomplete picture of the review process. Second, we are interested in separating situations where revisions in targets reflect an ex post rationalization (i.e., some months into the program actual data for part of the year become available and targets may be simply revised to conform with actual developments) from situations where revisions represent a genuine forward-looking policy response in the face of changing economic circumstances. Revisions at year T are often too close to the end of the first program year to be able to separate the two effects.

[5]A Chi-square test rejects the null hypothesis of no association between the two distributions. Chi-square = 23.08; df = 0.4; p-value = 0.0001.

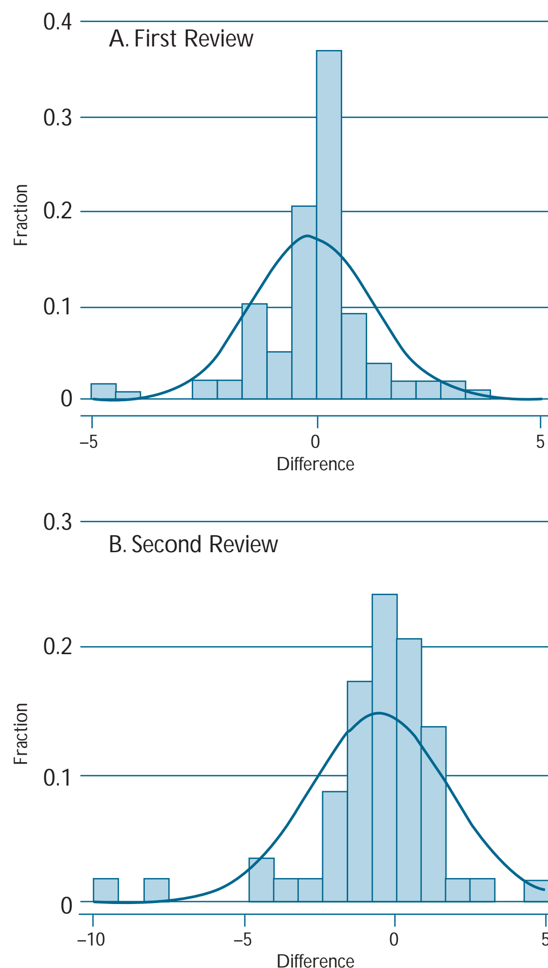
GDP growth was adjusted downward and for those where it was adjusted upward. The results show that when growth is revised downward by 1 percentage point, fiscal targets (as percent of GDP) are on average revised downward by ⅓ of 1 percent of GDP. However, no statistically significant impact was found on fiscal targets when growth was revised upward.[6]

These results suggest that program targets do respond to changes in growth outlooks as the program unfolds but the response tends to be asymmetric. Fiscal targets are revised downward when growth is below expectations, but they are less often revised upward when growth turns out to be higher than originally projected.

Rationale for revisions in fiscal targets

The rationale for the mid-course revisions of fiscal targets was further examined on the basis of the 15 case studies used in this evaluation. Table 4.8 shows the revision in targets for these case studies. We selected 11 cases in which the fiscal deficit target was adjusted upward or downward by more than 1 percentage point of GDP between the original program request and the first review (Algeria, Ecuador, the Philippines, Romania, Uruguay, and Venezuela programs), or between the first review and the second review (Bulgaria, Jordan, the Philippines, Senegal, and Tanzania programs).[7] Three of these cases (Ecuador, Bulgaria, and Venezuela) were examples where the fiscal target was actually tightened, whereas in the other eight it was relaxed.

Figure 4.1. Distribution of Programs According to Differences in Fiscal Adjustment Between Original Targets and Reviews
(In percent of GDP)



Sources: MONA and WEO databases.
Note: Negative (positive) values correspond to situations in which the review has lowered (raised) the targeted improvement in the fiscal balance.

The reasons given in the review documents for the revision in fiscal targets are summarized in Table 4.9 classified into two groups: those with little or no explanation, and those that provide some justification for the revisions. We find that out of the 11 cases in which the fiscal target was revised by more than 1 percent of GDP, program documents provided some justification for the new target in 7 cases with little or no justification in the other 4.

Two patterns emerge that are worth reporting:

(1) When fiscal performance by the time of the review was weaker than projected, program documents did not clearly analyze and try to separate

---

[6]This result is based on the following regression framework:

$$\Delta GBAL_{T+1}^{REVIEW1} - \Delta GBAL_{T+1}^{BOARD} = \alpha + \beta * [\Delta GDP_{T+1}^{REVIEW1} - \Delta GDP_{T+1}^{BOARD}]$$

Where $\Delta GBAL_{T+1}^{BOARD}$ is the original targeted change in the fiscal balance from T–1 to T+1; $\Delta GBAL_{T+1}^{REVIEW1}$ is the targeted change in the fiscal balance at the time of the first review; ΔGDP_{T+1}^{BOARD} is the original envisaged rate of GDP growth for T+1; and $\Delta GDP_{T+1}^{REVIEW1}$ is the envisaged rate of GDP growth for T+1 as projected at the time of the first review. In the baseline regression model, $\beta = 0.23$ (statistically significant at the 99 percent or better level of confidence). This suggests that for every 1 percentage point that GDP is revised downward, the targeted fiscal adjustment is reduced about ¼ of 1 percent. In principle, the regression coefficient could also be interpreted in the opposite direction, namely, that for every 1 percent that GDP is revised upward, the targeted fiscal adjustment increases by ¼ of 1 percent. However, further analysis does not warrant this conclusion. When we ran the regression separately for the cases in which growth was revised upward and the cases in which it was revised downward, the results showed that the unstandardized beta coefficient was about 0.32 and statistically significant when growth was revised downward, but close to zero and insignificant when growth was revised upward.

[7]We focus on cases of relatively large revision in the initial fiscal target where the need to explain why the new target is needed is presumably more relevant.

Table 4.7. Distribution of Programs According to Revisions in Growth and Fiscal Balances
(First Review)

Revisions in GDP Growth \ Revisions in Fiscal Balances	No change	Upward Revision	Downward Revision	Row Total
No change	**14**	10	6	30
Upward revision	1	**14**	14	29
Downward revision	4	13	**21**	38
Column total	19	37	41	97

Source: MONA database.

Table 4.8. Revisions in Fiscal Balance Targets for T+1
(In percent of GDP)

	Fiscal Balance T–1	Original Program	First Review	Second Review
Algeria SBA 1994[1]	–8.6	3.3	1.2	N/A
Bulgaria EFF 1998[1]	–2.5	–2.0	–2.8	–1.5
Costa Rica SBA 1995	–5.1	–1.1	–1.1	N/A
Ecuador SBA 2000[2]	–5.8	–3.9	–2.8	N/A
Egypt SBA 1996	–1.3	0.2	–0.6	–0.9
Jordan EFF 1999[3]	–9.5	–5.5	–5.5	–7.5
Pakistan SBA 2000	–6.0	–5.3	–5.3	–5.3
Peru EFF 1996	–2.8	–0.7	–0.7	–0.8
Philippines SBA 1998[1]	–0.7	–0.1	–1.6	–2.8
Romania SBA 1999[1]	–5.0	–1.1	–3.4	N/A
Senegal PRGF 1998[1, 3]	–2.0	–1.0	–1.0	–4.4
Tanzania ESAF 1996[1]	–4.4	–2.0	–2.0	–3.4
Ukraine EFF 1998	–5.4	–1.0	–1.0	–1.3
Uruguay SBA 2000[1]	–4.2	–1.2	–2.6	–3.3
Venezuela SBA 1996[2]	–3.2	–3.3	–0.4	N/A

Source: Program documents.
Note: N/A = not applicable.
[1]Cases of revision between the original program target and the first review or between the original program and the second review, more than 1 percent of GDP.
[2]For Ecuador and Venezuela we examined original program projections and reviews for year T due to insufficient data for year T+1. For all other programs, figures refer to changes in the government fiscal balance from T–1 to T+1.
[3]Fiscal balance excluding grants.

what part of this weaker performance was due to weak policy implementation and what part to factors outside the control of governments—for example, lower growth than expected, higher interest payments, terms of trade shocks, etc. Furthermore, there is a general tendency to emphasize the role of factors outside the government's control. However, as we show later, insufficient progress in structural reforms in the fiscal area is an important factor behind shortfalls in fiscal adjustment in IMF-supported programs.

(2) Review documents tend to be more backward-looking than forward-looking; they typically elaborate why past fiscal developments call for revisions, but not much is said why the new targets are appropriate given the overall objectives of the program.[8] Only the Philippines program, and to a lesser extent the Uruguay program, provided sufficient forward-looking analysis or the need for revising fiscal targets. In the case of the Philippines program, the review documents for the first and second reviews contain a comprehensive analysis

[8]This is certainly understandable in those cases in which the review takes place very close to the end of year T+1 (as there is little time to change course in fiscal policy). But when the distance between the time of the review and the end of T+1 is sufficiently large (e.g., more than six months), review documents should explain why the new target is consistent with the overall objectives of the program.

Table 4.9. Summary of Reasons for the Revised Fiscal Balance Target

	Little or No Discussion	Relatively Clear Discussion
Algeria SBA 1994 (First review)	No explanation for the downward revision in the fiscal target for year $T+1$.	
Ecuador SBA 2000 (First review)		Higher growth and higher oil prices than expected led to better revenue performance. Lower expenditures due to cuts in investment spending.
Jordan EFF 1999 (Second review)		Prudent fiscal stance is central to achieving macroeconomic objectives. Need to diminish high public debt burden. Mindful that unduly rapid fiscal adjustment would have a negative effect on growth and employment.
Philippines SBA 1998 (First review)		Loosening of the fiscal target to accommodate the effects of sharply lower GDP growth and somewhat higher social spending than originally envisaged.
Romania SBA 1999 (First review)	Review document contains a comprehensive analysis of fiscal policy but does not provide a good sense of why the fiscal deficit target was revised downward.	
Uruguay SBA 2000 (First review)		Loosening of fiscal target owing to expected revenue shortfalls associated with a large output gap relative to potential GDP. Adhering to the original fiscal deficit target, the review argues, would imply large tax increases that would make the resumption of growth more difficult.
Venezuela SBA 1996 (First review)		Review document explains that higher-than-expected oil prices and lower-than-expected interest payments justify the much larger improvement in the fiscal balance than originally expected.
Bulgaria EFF 1998 (Second review)		Fiscal target is tightened to anticipate a worse-than-expected external position. Stronger revenues than initially expected also played a key role.
Philippines SBA 1998 (Second review)		Fiscal targets are relaxed to accommodate the effects of weaker economic conditions and also to permit higher social expenditures.
Senegal PRGF 1998 (Second review)	No explanation for the revision in the fiscal target.	
Tanzania ESAF 1996 (Second review)	No explanation for the revision in the fiscal target.	

Source: Program documents.

of the fiscal stance (EBS/98/172). They provide an assessment of the past and an analysis of the new fiscal balance target for $T+1$. The report justifies the relaxation of the fiscal target in terms of lower growth than originally envisaged and the need to accommodate higher social spending. In the case of Uruguay the review document argues that, given the weak revenue performance by the time of the review, sticking to the original fiscal target would have a contractionary impact on output.

Table 4.10. Envisaged and Actual Fiscal Adjustment in Nine IMF-Supported Programs
(Ranked by the magnitude of the fiscal shortfall)

Program	Initial Conditions	Fiscal Adjustment (T–1 to T+1)			Decomposition of the Fiscal Shortfall[1]	
	Fiscal balance (T–1)	Envisaged	Actual	Difference	Revenue component	Expenditure component
		(In percent of GDP)			*(In percent of the overall adjustment)*	
Cases of fiscal underperformance						
Algeria	–8.6	11.9	7.2	–4.7	–25	–75
Philippines	–0.7	0.8	–3.5	–4.3	–93	–7
Tanzania	–4.4	2.7	–0.7	–2.0	–90	–10
Romania	–5.0	3.9	1.0	–2.9	–175	+75
Costa Rica	–5.1	4.0	1.1	–2.9	–155	+55
Uruguay	–4.2	3.0	0.0	–3.0	–7	–93
Jordan	–9.5	4.0	0.6	–3.4	–166	+66
Ukraine	–5.4	4.4	3.1	–1.3	–115	+15
Egypt	–1.3	1.5	0.3	–1.2	–75	–25

Source: Program documents.
[1]The percentage contribution of revenue plus expenditure shortfalls add up to the 100 percent shortfall in the fiscal adjustment. Negative (positive) values show that revenues or expenditures adjusted less (or more) than was projected. When the revenue component adds up to more than minus 100 percent, it means revenue shortfalls were larger than the total fiscal adjustment shortfall—expenditures then adjusting more than projected.

What Accounts for Large Fiscal Underperformance? Evidence from the Smaller Sample

The analysis of differences between envisaged and actual fiscal adjustment in the large sample of programs can be supplemented by evidence from the smaller sample of 15 from which we extract the 9 cases of fiscal underperformance (Table 4.10). Specifically, we focus on the 7 cases of large fiscal shortfall, that is, where actual adjustment was 2 or more percentage points of GDP less than envisaged. These 7 cases of large fiscal shortfall comprise 2 (Algeria SBA 1994 and Uruguay SBA 2000) dominated by expenditure overruns and 5 (Costa Rica SBA 1995, Jordan EFF 1999, the Philippines SBA 1998, Romania SBA 1999, and Tanzania ESAF 1996) mainly caused by revenue shortfalls.

In the case of the Uruguay and Algeria programs, shortfalls in the expenditure/GDP ratio relative to targets cannot be attributed to weaknesses in implementing the program. In the case of Uruguay, nominal spending was in fact within the agreed ceiling, with the shortfall reflecting a significantly lower growth performance than projected. (In fact, the program projected recovery of growth while in reality growth was negative.) In the Algeria program, nominal spending was indeed higher than envisaged. However this reflected unexpected shocks: specifically, spending in the wake of an earthquake, and higher-than-expected outlays to protect public safety in response to heightened security concerns.

In contrast, the revenue shortfalls seem to be associated with weak implementation as outlined below. Table 4.11 compares the shortfall in envisaged GDP growth to the revenue shortfall in the five large revenue underperformers.

In most cases (except Costa Rica) growth improved during the program period. Despite this acceleration in growth, revenue ratios declined. This suggests that GDP growth played a limited role in accounting for the poor revenue performance. Neither could the large shortfall in revenue performance with respect to targets be explained by the observed shortfall in growth performance in the program period relative to projections. Indeed, revenue underperformance is about four times the growth underperformance. These magnitudes cannot be explained by typical revenue-GDP elasticities, which are normally around one.[9]

We can summarize the above findings as follows:

- For the programs with expenditure shortfalls, it is either unexpected shocks or the optimism in the envisaged GDP growth that explains expenditure overruns as a share of GDP.

[9]In some cases, growth may have been concentrated in lightly taxed sectors (e.g., agriculture and exports in the Philippines), a factor that may not have been anticipated in the original revenue projections. Nevertheless, if the shift to lightly taxed sectors was permanent, relatively painless policy action should have been feasible to restore the traditional share of taxes by taxing part of the unexpected growth, for example by rolling back exemptions.

Table 4.11. Comparing Growth and Revenue Underperformance

	Growth (In percent)		Δ Revenues (In percent of GDP)	Shortfall with Respect to Program Targets	
Program	T	T+1	T–1 to T+1	GDP growth[1] (Percentage points)	Revenue[2] (In percent of GDP in T+1)
Costa Rica	2.4	0.7	–2.3	–1.2	–4.5[3]
Jordan	3.1	4.0	–2.4	1.3	–3.0
Philippines	–0.6	3.4	–3.6	–2.1	–3.4
Romania	–2.3	1.6	–1.3	0.1	–5.2[3]
Tanzania	3.7	3.7	–0.8	–1.8	–4.5
Average	1.3	2.7	–2.1	–0.8	–4.1

Source: Program documents.
[1]Difference between the actual average growth in *T* and *T*+1 and the equivalent projected value.
[2]Difference between the actual revenue over GDP in *T*+1 and the equivalent projected value.
[3]Consistency of data may be compromised by data revisions in the GDP series after the original program request. The revenue shortfall after taking these revisions into account is still substantially large.

Table 4.12. Revenue-Related Structural Reform Measures in Selected Programs with Large Revenue Shortfalls

Program		Implementation
Philippines	• Suspend all tax subsidies of national government agencies.	Partial progress.
	• Strengthen tax administration.	Partial/poor progress.
	• Continue comprehensive tax policy reform.	Slow progress.
	• Reorganize Large Tax Payer Division.	Partial progress.
Tanzania	• Reduce tax evasion through: (1) Harmonization of import taxes between the mainland of Tanzania and Zanzibar; (2) audit of bonded warehouses and establishment of a monitoring system prior to computerization.	Done.
	• VAT legislation to be passed by parliament and for *T*+1 administrative measures to support VAT introduction.	Done with delay.
Romania	• Increase excise taxes, property taxes, and social security contributions.	Partial progress.
	• Eliminate tax exemptions.	Done.
	• Delay tax decreases approved during 1998.	Implemented with delay.
	• Collect tax arrears.	Not done.
Jordan	• Reduction in the maximum import tariff to 30 percent.	Done.
Costa Rica	No structural benchmarks related to revenue but there was a PC on the net borrowing requirements of the nonfinancial public sector that incorporated an anticipated 3 percentage point of GDP increase in taxes from:	A waiver was required for the PC due, inter alia, to delays in adopting tax measures in 1995.
	• an increase in the sales tax rate from 10 percent to 15 percent for 18 months before falling back to 13 percent;	
	• a new export tax structure for coffee to capture some of the windfall from higher prices;	
	• a 1 percent tax on gross assets of corporations; a consumption tax on petroleum products; and	
	• the unification of the tax rate on company profits.	

Source: Program documents.

- In programs with significant revenue shortfalls with respect to targets, neither actual growth performance nor growth optimism can explain these shortfalls.

Optimism in growth projections cannot, therefore, explain the large underperformance of revenue. Instead, underperformance must be related to other factors such as structural reforms. Either reforms were implemented less rapidly than envisaged or staff overestimated the impact of these reforms. Indeed, Table 4.12 suggests that underperformance was mainly the result of insufficient progress in revenue-enhancing structural reforms. Cases of large revenue shortfalls were mostly the result of poor implementation of structural reforms envisaged in the program (e.g., the Costa Rica, Jordan, Philippines, and Romania programs), or the implementation of reforms likely to have an impact on revenue only over the medium term (e.g., the Tanzania program that envisaged preparatory steps for the implementation of VAT and parliamentary approval of associated legislation).

Chapter 5

Economic Recovery and Growth

The pace of economic recovery under an IMF-supported program is an important determinant of the impact of a program on welfare and is, therefore, the focus of considerable attention. To the extent growth performance is significantly below program projections, it may signal shortcomings in program design, including the fiscal stance. As pointed out in the introduction, one of the criticisms of program design in IMF-supported programs is that they impose unduly tight fiscal policies leading to adverse effects upon economic recovery. In this chapter, we examine these issues using the large sample of programs studied for this evaluation.

Economic Recovery in the Program Period: Outcomes and Expectations

Table 5.1 summarizes the short-term growth experience in 159 IMF-supported programs.[1] It presents average annual growth rates actually achieved during preprogram and program years for the whole sample as well as for the subgroups used for this study. We note that comparisons between preprogram and postprogram growth recovery should not be understood as indicating the impact of the program, but only as a description of what happened. The impact of IMF-supported programs on growth can only be determined by comparing actual outcomes with the counterfactual of what would have happened to economic performance without a program. There is now an extensive, albeit inconclusive, literature on the topic but this area goes beyond the limits of this evaluation.[2]

The following are the main features that emerge from the data:

- For all programs taken together, the average GDP growth rate achieved in the first program year *T* improved upon the level in *T*–1, and then improved further in *T*+1 when it actually surpassed the average of the preprogram decade.
- The same pattern is discerned in the two subgroups consisting of ESAF/PRGF and SBA/EFF transition cases. However, the SBA/EFF nontransition group shows a somewhat different behavior, with the average growth rate decelerating sharply from 2.4 percent in *T*–1 to 0.9 percent in *T*.
- The SBA/EFF nontransition subgroup itself consists of two very different types of programs. There are 10 programs in this group which relate to so-called capital account crises while the others relate to more conventional balance of payments problems.[3] The capital account crisis cases experienced a collapse in output with average GDP growth falling sharply to –5.0 percent in *T* compared with 2.9 percent in *T*–1. This was followed by a recovery in *T*+1, which almost offset the decline in the previous year. The other 51 programs in this category show only marginal deceleration, with growth decelerating from 2.3 percent in *T*–1 to 2.1 percent in *T*, followed by a respectable acceleration to 3.5 percent in *T*+1.

[1]This is the maximum number of arrangements for which a comparable series of data on projected and actual growth could be obtained from the MONA and WEO databases. This sample is larger than the sample of 133 arrangements used previously to compare developments in fiscal balances.

[2]The literature follows the original study by Goldstein and Montiel (1985), which tried to isolate the impact of an IMF-supported program by the Generalized Evaluation Estimator (GEE), and the subsequent study by Khan (1990). The GEE attempts to provide a measure of the policies that would have prevailed in the absence of an IMF-supported program. Although some earlier studies have shown no impact or a negative impact of IMF-supported programs on growth, the results of this line of research have been rather sensitive to model specification and the choice of variables included in the analysis. For a general review of the literature on this topic, see Joyce (2002) and Haque and Khan (1998).

[3]The distinction between capital account crises and other more conventional balance of payments crises that have their origin in the current account is now well established, though it is not always as sharp as it sometimes appears because even conventional current account crises may generate capital account feedback effects. The 10 programs identified for inclusion in this group are the eight IMF-supported programs identified in Ghosh and others (2002) (Argentina 1995; Brazil 1998; Indonesia 1997; Korea 1997; Mexico 1995; the Philippines 1997; Thailand 1997; Turkey 1994) plus Turkey 1999 and Argentina 2000.

Table 5.1. Experience with GDP Growth Prior to and During Program Periods: Annual GDP Growth
(In percent)

	Number of Programs	Trend in the Prior Decade	*T*–1	*T*	*T*+1
All programs	159	1.6	1.4	2.2	3.8
ESAF/PRGF	64	1.7	2.8	4.4	4.3
SBA/EFF (transition countries)	34	–2.1	–3.3	0.4	3.0
SBA/EFF (nontransition countries)	61	3.6	2.4	0.9	3.7
Of which					
Capital account crisis cases	10	4.8	2.9	–5.0	4.7
Other programs	51	3.4	2.3	2.1	3.5

Source: WEO database.

The average growth rates presented in Table 5.1 suggest that the perception that IMF-supported programs are associated with strongly negative effects on growth is not well founded, except in the case of capital account crises.[4] However, averages can be misleading because of variations around the mean, so we have also examined the distribution of programs and identified the percentage of programs which show a deceleration in growth compared with *T*–1 and those that show negative growth. Two different time horizons are used for the program period, a one-year horizon *T* and a two-year horizon covering *T* and *T*+1. The two-year horizon is perhaps more relevant since many programs only commence in the middle of *T*. The results are presented in Table 5.2.

Although the average growth rate of all programs did not decelerate (see Table 5.1), it is clear that a substantial percentage of programs in all subgroups experienced a deceleration in growth not only over a one-year but also over a two-year horizon. The number experiencing negative growth is much smaller and this phenomenon is concentrated in the group of transition and the capital account crisis cases. In the transition cases, the negative growth is actually a continuation of negative growth in the preprogram period (see Table 5.1). In the other two groups, ESAF/PRGF and SBA nontransition others, negative growth over a two-year horizon was experienced by only a small proportion (3 percent and 16 percent, respectively) of cases.

[4]Again, these results do not "prove," in any sense, that IMF-supported programs are good or bad for the recovery of growth. For example, because of mean-reversion phenomena (i.e., the tendency of an economy to revert to normal growth rates after a shock) it could be argued that growth would be expected to be stronger in any event in years *T* and *T*+1, merely because the impact of the adverse shock that caused the country to seek IMF support would tend automatically to dissipate as time passes.

The actual recovery and short-term growth performance in the postprogram period also needs to be compared with GDP growth projections in programs. This comparison is important because the public perception of the success of programs is often assessed not just in terms of the actual outcomes but in terms of achievement relative to the growth targets. Furthermore, large shortfalls in growth relative to projections can generate consequential problems because fiscal targets built into programs may become inappropriate.

To compare projections with actuals we use the cumulative growth over *T* and *T*+1 as the basis for comparison (Table 5.3). The main conclusions are the following:

- Actual growth fell short of projected growth over the two-year period and the average shortfall for all programs amounted to 1.5 percentage points. Except for the subgroup of transition countries, where the actual two-year achievement is basically the same as projected, all the other groups show underperformance on average.
- The shortfall in the case of ESAF/PRGF is 1.5 percentage points, the same as for all programs. The shortfall in the SBA/EFF nontransition cases in turn reflects divergent behavior in the two subdivisions within this group. There was a massive underperformance of 6.4 percent in the case of the 10 capital account crisis cases. The other programs in the subgroup show only a relatively modest shortfall of 1.6 percent of GDP, which is close to the average of 1.5 percent for ESAF/PRGF.[5]

[5]This finding is also consistent with a recent study by Musso and Phillips (2002). The study found a tendency toward growth optimism in programs involving large access to IMF resources, those usually associated with crisis situations and large capital flow reversals.

Table 5.2. Programs Showing Deceleration or Negative Growth

	One-Year Horizon		Two-Year Horizon[1]	
	Percentage of programs showing		Percentage of programs showing	
	Deceleration	Negative growth	Deceleration	Negative growth
All programs	42	36	36	18
ESAF/PRGF	44	9	39	3
SBA/EFF (transition countries)	20	48	10	40
SBA/EFF (nontransition countries)				
Of which				
Capital account crisis cases	80	90	80	40
Other programs	47	23	41	16

Source: WEO database.

[1]Programs are classified as indicating deceleration or negative growth on the basis of the annual average growth rates in the two-year period T and T+1. A negative average growth rate over two years means GDP in T+1 was lower than in T–1.

Table 5.3. Envisaged and Actual Two-Year Cumulative Growth Rates over T and T+1
(In percent)

	Projected Cumulative Growth	Actual Cumulative Growth	Shortfall (Actual – Envisaged)
All programs	7.7	6.2	–1.5[1]
ESAF/PRGF	10.5	9.0	–1.5[1]
SBA/EFF (transition countries)	3.5	3.6	0.1
Of which			
Capital account crisis cases	5.8	–0.5	–6.4[1]
Other programs	7.3	5.7	–1.6[1]

Sources: MONA and WEO databases.

[1]Difference statistically significant at the 95 percent or better confidence level.

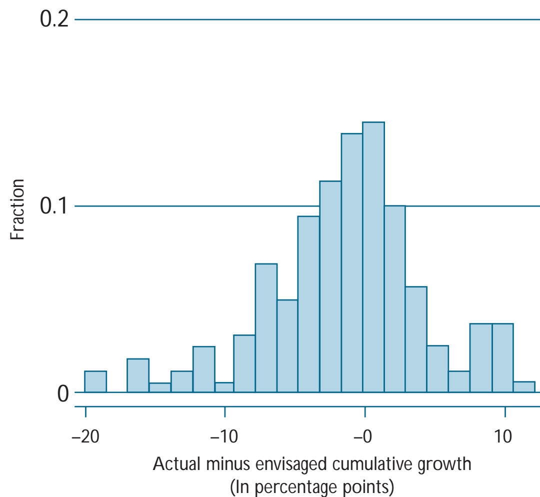
Since there is considerable variation around the means reported in Table 5.3, it is useful to look at the distribution of programs according to the differences between actual and envisaged cumulative growth (Figure 5.1). About 60 percent of the cases show a shortfall. In about 25 percent of programs, the shortfall in cumulative growth over the two-year horizon exceeds 4 percentage points.

Table 5.4 also suggests that the degree of optimism about growth in T+1 depends upon what has happened in T. For all programs, the growth rate projected for T+1 was too optimistic by 1 percentage point. However, for those programs where growth was negative in year T (one-quarter of the overall sample), the growth projected for T+1 was subject to greater overoptimism (double the actual growth). Program projections of growth tend to build in greater optimism about recovery when starting from an adverse situation, probably reflecting an understandable expectation of reversal to normality.

Even more striking is that programs are reluctant to project slowdown in growth from T to T+1, let alone to project negative growth. Only 18 percent of programs projected a slowdown in growth, whereas this happened in almost 40 percent of cases. Programs seldom project negative growth, although in reality it happens in about 13 percent of cases. Programs tend to underpredict significantly more situations of adverse output developments than situations of favorable output developments.[6] This tendency must be seen in the context of the fact that program

[6]We found that programs forecast 1.3 percent of cases as having negative growth in T+1, while in reality this happens in 13 percent of cases. On the other hand, programs forecast 5 percent of cases to have growth larger than two times mean growth (a symmetrical deviation from the mean) while in reality this happened in 11 percent of cases. Thus, programs systematically underpredict negative output developments relative to favorable developments.

Figure 5.1. Distribution of Programs According to Differences Between Actual and Envisaged Cumulative Growth over a Two-Year Period (*T* and *T*+1)



Sources: MONA and WEO databases.

Table 5.4. Indicators of Growth Optimism for *T*+1

A. Growth rate at T+1
(In percent)

	Envisaged	Actual
All arrangements	4.6	3.6
Arrangements where growth was negative in year *T*	3.4	1.7

B. Frequency of cases
(In percent)

	Envisaged	Actual
Percentage of cases where growth rates are reduced from *T* to *T*+1	17.8	39.5
Percentage of cases where growth in *T*+1 is negative	1.3	12.6

Sources: MONA and WEO databases.

projections are not just the outcome of technical analysis but are negotiated outcomes and there are strong compulsions to present as optimistic a picture as possible. Nevertheless, it does suggest that the reluctance of programs to "call a downturn" means that the appropriate fiscal stance in such circumstances is not addressed in the original program design.

Optimism in Projecting Private Demand and Investment

There are many reasons why growth outcomes during the recovery phase might differ from projections. These include (1) exogenous factors turning out to be different from what was expected; (2) policies on which the growth projection was based may not be implemented effectively; (3) the projections may have been based on an inadequate understanding of the determinants of short-term growth leading to an inadequate design of policies; and finally (4) acceptance of an overoptimistic projection as an outcome of the program negotiation process. These factors must have operated to different degrees in different programs and it is beyond the scope of this evaluation to go into all these issues. However, there is one factor which may explain some of the optimism about growth in many cases, which can be examined with the available data, and this relates to the tendency to be overoptimistic in projecting investment, especially private investment.

Crisis situations are typically disruptive and introduce uncertainty about economic outcomes that can be expected to have a temporary negative effect on private investment. The rate of economic recovery may depend significantly on the pace at which investment activity goes back to normal. Unfortunately, the MONA database does not contain data on projected private investment in programs. However, it contains information on projected total investment rates and this can be used to examine the extent of overoptimism regarding total investment and its possible relationship with growth shortfalls.

Earlier IMF staff studies have documented that IMF-supported programs typically overestimate the speed with which investment will recover.[7] Table 5.5, which presents available information on actual and projected investment rates for the large sample and for the individual subgroups, confirms that there was overoptimism on average for all programs and the extent of optimism increases from *T* to *T*+1.[8]

[7]Goldsbrough and others (1996). Moreover, there is a large theoretical and empirical literature suggesting that a lagged response of private investment should be expected following a period of adjustment. See, for example, Dixit and Pindyck (1994) and Serven and Solimano (1994).

[8]Regression results (not shown) also suggest a strong and statistically significant link between the projected acceleration of growth in programs and the projected increase in investment rates.

Table 5.5. Investment Projections and Actuals Under IMF-Supported Programs, 1993–2001

(Annual investment in percent of GDP)

		T			*T*+1		
	T–1	Projected	Actual	Difference[1]	Projected	Actual	Difference[1]
All programs	20.6	21.0	20.7	–0.3	22.0	21.2	–0.8
ESAF/PRGF	18.6	18.6	19.4	0.8	19.8	21.0	1.2
SBA/EFF (nontransition countries)	22.6	23.3	22.4	–0.9	23.9	21.7	–2.2[2]
Of which							
Noncapital account	22.2	23.1	22.6	–0.5	23.8	21.6	–2.2[2]
Capital account	24.5	24.1	21.2	–2.9[3]	24.6	22.5	–2.1
SBA/EFF (transition countries)	21.1	21.6	20.1	–1.5[3]	23.0	20.7	–2.3[2]

Source: MONA and WEO databases.
[1]Difference is actual minus projected.
[2]Significant at the 95 percent confidence level.
[3]Significant at the 90 percent confidence level.

There are interesting differences in investment behavior among the various subgroups and its relationship with growth outcomes.

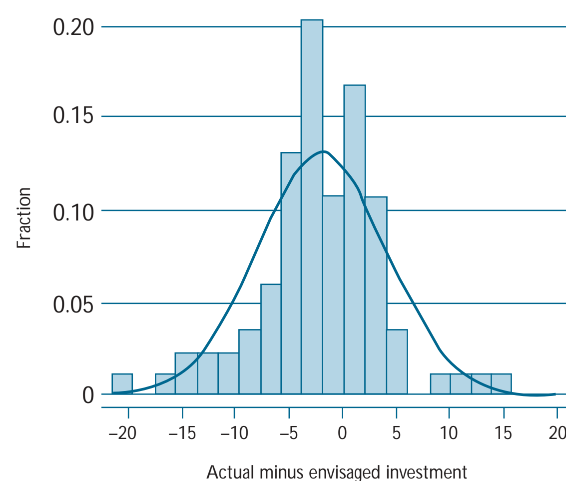
- In ESAF/PRGF programs actual investment rates are slightly higher than program projections. Thus, the shortfall in growth experienced by this subgroup cannot be attributed to investment shortfalls.
- In SBA/EFF transition cases, actual investment rates fall short of projected levels by 1.5 percentage points in *T* and 2.3 percentage points in *T*+1, but the GDP growth rates achieved are very close to projections. This suggests that other positive factors, possibly the pace and impact of structural change in these countries, must have been stronger than expected and offset the negative impact of investment shortfalls.
- The subgroup of SBA/EFF nontransition cases shows significant investment shortfalls and as seen in Table 5.3, this group also showed cumulative growth shortfalls.

As in other comparisons based on group averages, it is useful to look at the extent of variation. Figure 5.2 shows the distribution of the differences between actual and projected investment rates for *T*+1 for SBA/EFF programs. Investment rates were below projections in about two-thirds of programs. In about one-quarter of programs, investment rates were 5 percentage points of GDP or more below projections. The proportion of cases where programs predicted a decline in investment rates between *T*–1 and *T*+1 was also seriously underestimated. Programs projected a decline in 25 percent of cases, while in reality investment rates declined in 50 percent of cases.

To summarize, there is evidence of generalized optimism in programs regarding the prospects for investment. That optimism becomes significantly magnified in the context of projecting recoveries from adverse initial conditions, and/or projecting the possibility of deterioration in performance. The problem is especially important in the case of SBA/EFF nontransition cases, and within this group, the set of capital account crisis episodes appears to be worse affected. The specific assumptions that may be behind this optimism in private spending and investment projections and how these assumptions are linked to program instruments are critical to assess the appropriateness of the fiscal stance of programs. This is discussed next.

Figure 5.2. SBA and EFF Programs According to Differences Between Actual and Envisaged Investment Rates for *T*+1

(In percent of GDP)



Sources: MONA and WEO databases.

Is the Fiscal Stance in IMF-Supported Programs Unnecessarily Contractionary?

The issue of whether IMF-supported programs suffer from an unnecessarily contractionary fiscal stance has attracted special attention following some of the recent capital account crisis cases—notably Korea and Indonesia, both of which experienced large output declines and an increase in unemployment. Critics have argued that the fiscal adjustment proposed in those cases was inappropriate and may have even contributed to worsening the situation. The specific cases of Korea and Indonesia have already been examined in detail in an earlier IEO report and are not discussed individually in the present evaluation.[9] In this report, we consider what light can be shed on this issue from the broader cross-country evidence studied for this evaluation. For this purpose, we focus on SBA/EFF programs in nontransition countries as this is the group where the problem of a contractionary effect is perhaps most relevant.

Table 5.6 presents some of the critical macroeconomic data distinguishing between the capital account crisis cases (Panel A) and noncapital account crisis cases (Panel B). The following features are relevant to our evaluation.

- The capital account crisis cases experienced a severe output contraction in year *T*, resulting in a massive underperformance in output relative to expectations. The noncapital account crisis cases do not show an output contraction on average, but they do show a shortfall in growth compared with projections, especially in *T*+1.
- Both groups show an underperformance in investment rates relative to expectations with the phenomenon being more marked in the case of capital account crisis cases.
- Both groups show an underperformance on the fiscal side with fiscal deficits significantly higher than program targets. Again, the phenomenon is more marked in the case of capital account crisis cases, reflecting the decline in GDP in these cases and the asymmetric response of revenues and expenditures.
- Both groups also show overperformance on the external side, in that the current account deficit was reduced much more than programmed. This is particularly so in the capital account crisis cases where the current account adjustment on average was 4.8 percent of GDP higher than programmed in year *T* and 2.6 percent of GDP higher in *T*+1. The corresponding numbers for the noncapital account cases are 0.9 percent and 1.2 percent of GDP, respectively.

The experience of the noncapital account crisis cases appears to be a milder form of the experience of the crisis cases, with the problem surfacing not in a decline in output but in a shortfall in growth performance in *T*+1.

The fact that both output and investment were below programmed levels raises the possibility that these may be classic cases of Keynesian lack of effective demand, in which higher levels of output could have been achieved if fiscal policy in the short run had been less contractionary. This perception is reinforced by the fact that the current account deficit overcorrected compared to projections, even though the fiscal targets originally projected in the program were not achieved. This can be viewed as suggesting that the original fiscal deficit targets were excessively tight and a more relaxed fiscal stance might have allowed higher levels of output and employment. Of course, the current account deficit could be expected to widen in this situation, but since the data show overcorrection in this dimension, it can be argued that there was room for some deterioration while leaving the deficit within the financeable range.

The emphasis on tightening fiscal policy could be traced to unrealistic assumptions about the pace at which private investment demand will recover following the crisis. Programs typically assume rapid recovery, and therefore tend to push for greater fiscal adjustment to make room for private investment, whereas a more realistic recognition of the negative impact of crises on investor expectations would call for a more relaxed fiscal stance.

It can be argued that a more expansionary policy may not have been feasible if external financing was not available to finance the resulting increase in the current account deficit. However this does not seem to be the case in a number of the programs we have examined. One-fourth of the SBA/EFF programs in nontransition cases showed overperformance not only in the current account, but also in the buildup of reserves. External financing does not seem to have been a constraint in these cases and a less contractionary fiscal stance could have been more appropriate.

This essentially Keynesian argument focuses exclusively on the role of fiscal adjustment as a factor affecting aggregate demand. However, as pointed out in Chapter 2, "Relevant Considerations in Determining the Fiscal Stance," this is only one of the factors relevant in determining the fiscal stance. Emerging market countries relying on international

[9]See IEO (2003).

Table 5.6. Macroeconomic Balances in Stand-By and EFF Arrangements in Nontransition Countries[1]
(In percent of GDP)

A. Capital account crisis cases

	T–1	*T*		*T*+1	
N = 10	Actual	Envisaged	Actual	Envisaged	Actual
Current account	–3.4	–2.4	2.4	–2.4	0.2
Government balance	–3.3	–1.8	–4.3	–1.6	–3.7
Total investment	24.5	24.1	21.2	24.6	22.5
Private sector balance	–0.1	–0.6	6.7	–0.8	3.9
GDP growth (in percent)	2.9	1.6	–5.0	4.1	4.7

B. Noncapital account crisis countries[2]

	T–1	*T*		*T*+1	
N = 45	Actual	Envisaged	Actual	Envisaged	Actual
Current account	–3.1	–3.1	–2.2	–3.4	–2.2
Government balance	–4.0	–2.4	–3.2	–1.4	–3.5
Total investment	22.2	23.1	22.6	23.8	21.6
Private sector balance	0.9	–0.7	1.0	–2.0	1.3
GDP growth (in percent)	2.1	2.5	2.2	4.5	3.2

Sources: MONA and WEO databases.

[1]All differences between envisaged and actual values are statistically significant (with the exception of growth in *T*+1 in capital account crisis cases; and growth in investment in *T* for noncapital account cases).

[2]The average growth figures differ slightly from those in Table 5.3 because the Lesotho SBA programs (1994/1995/1997) and the Republic of Congo SBA program (1994) were excluded due to problems in the reliability of the current account data.

financial markets also have to consider the impact of the fiscal stance adopted in times of crises on market confidence and therefore the availability of external finance. Advocates of a tighter fiscal stance can legitimately argue that, in situations where debt sustainability is an issue, it may be necessary to accept a larger dose of fiscal adjustment to reassure markets and ensure revival of confidence, even though a more relaxed stance may be justifiable on countercyclical grounds. In this view, the benefits of countercyclical fiscal policy can only be enjoyed in circumstances where the underlying fiscal situation is sound and markets recognize that the relaxed fiscal stance reflects a temporary resort to automatic stabilizers, and not simply an unwillingness to take difficult decisions.

It is difficult to determine the extent to which the fiscal stance adopted in the various programs studied was the result of a conscious decision to send the right market signals and whether the scale of the adjustment proposed was appropriate under the circumstances. As pointed out above, our evaluation finds that program documents provide little analysis of the rationale for fiscal adjustment and its link with the recovery of private sector activity and growth. A clearer statement of the rationale would add to transparency by promoting better understanding of the different considerations involved in each case, with a fuller consideration of the underlying assumptions. It would also help to determine the degree of flexibility that must be shown at the time of program reviews. For example, there is a clear case for allowing flexibility in adjusting the fiscal deficit in the event that assumptions about investment demand prove overoptimistic. As pointed out above, IMF programs do show considerable flexibility in practice in revising fiscal targets, but the rationale for the revisions is often left unclear. This has the disadvantage that adjustments that are perfectly justified on grounds of automatic stabilization may be seen as a forced response to nonperformance, a perception that can undermine the very confidence which the program seeks to restore.

This evaluation recommends a series of practices in the design of future programs. Program documentation should explicitly discuss how the projected economic recovery is linked to assumptions on how private demand will respond to the impact of the program on confidence. This is critical for the discussion of the fiscal stance of the program. A tight fiscal stance is appropriate under the expectation of significant positive shifts in investment demand, thus creating room for this buoyant investment demand to be financed. However, the original fiscal stance may need to be modified if the same economic recovery is to be achieved with a less buoyant recovery of pri-

vate demand. In this case, it may be appropriate to include a stronger countercyclical element. These discussions may be particularly critical when private demand has initially collapsed as a result of a crisis situation at the outset of a program. A more careful identification of these links will provide a more coherent framework for sensitivity analysis. It will help to identify the critical assumptions and alert the staff early in the process on what needs to be monitored as the program unfolds.[10]

When staff decides to loosen fiscal policy to compensate for shortfalls in private demand, it should clearly evaluate the costs and benefits of doing so via the expenditure or revenue side. The mere fact that revenues may be lagging should not be an argument to accept a shortfall as the preferred way of temporarily widening the deficit. There may be good equity and solid reason to allow certain social expenditures momentarily to increase.

[10]In some instances, staff has also used independent output forecasts from academics or market analysts to complement program projections, for example, the recent staff reports of program reviews of the Brazil Stand-By Arrangement approved in 2002. This is a good practice that should be encouraged.

Chapter 6 Social Spending and Social Protection in IMF-Supported Programs

The impact of IMF-supported programs on the level of public spending in the social sectors has received a great deal of attention, with many critics voicing concern that these programs typically involve an unnecessary squeeze on social spending, with adverse effects on social welfare. We examine this issue in several ways. First, we analyze a set of concerns raised in the context of low-income countries—whether programs incorporate public spending levels and fiscal deficit targets based on overly conservative projections of concessional financing. Second, we examine cross-country data to assess what may have been the impact of IMF-supported programs on the level of public sector social spending. Third, we analyze program documents in the sample of 15 programs described earlier, to assess how program design has incorporated social spending and social concerns.

Has Donor Aid Been Underestimated?

Concerns have been raised that IMF-supported programs in low-income countries that depend on concessional financing may incorporate fiscal targets based on aid projections that "taper out" too quickly relative to what donors may be willing to provide. If true, such a tendency could also create a disincentive for donors to sustain their level of aid, even when programs remain on track.[1]

Some recent studies by IMF staff have argued in support of a cautious approach to projecting aid flows, mainly on the grounds that disbursements tend to be significantly less than commitments, and that even the so-called conservative projections in IMF-supported programs tend to overestimate actual aid flows.[2] These studies also point out that in the programs examined: (1) disbursements exceeded projected amounts in a minority of cases; (2) shortfalls relative to projections were more marked for program aid (compared to project aid); and (3) within program aid, grants (provided mainly by bilateral donors) had a smaller "prediction error" than concessional loans (a large part of which came from the World Bank and regional development banks).

One factor that may contribute to deviations between projections and outturns is compliance with conditionality. To the extent that the conditions attached to the disbursement schedule are not met, donors may withhold disbursements. For example, some donors link disbursements of their program aid to recipient countries' performance under IMF-supported programs. Thus, outturns in such cases are to some extent contingent on implementing policies in the program, and hence are endogenous. However, there is evidence that shortfalls occur even for programs that remain broadly on track.[3]

We have reexamined this issue by focusing on two questions:

(1) What is the extent of "tapering out" of projected donor flows between the initial and third year of the program? To address this question, we examined program projections in the MONA database for nearly 100 ESAF/PRGF arrangements approved during 1995–2001.[4]

(2) What are the differences between actual flows and projected levels of donor aid? To address this question, we undertook two exercises. One focused on revised projections for the first year of the program in each successive yearly arrangement under

[1]See, for example, Collier and Gunning (1999). The authors argue that the disincentive arises because programs usually do not allow additional aid (i.e., above the amount projected) to be spent, favoring instead the channeling of the extra amounts into increasing international reserves or paying down debt.

[2]See, for example, Bulíř and Hamann (2001) and Bulíř and Lane (2002).

[3]Bulíř and Hamann (2001) reported that countries with uninterrupted programs received, on average, about three-quarters of program aid commitments. Countries where programs were interrupted received only about one-third of program aid commitments.

[4]From November 1998, the three-annual-arrangement structure of the ESAF was replaced by a one three-year-arrangement structure. The comparison includes projections under both types of structure.

the typical three-year concessional program.[5] A second exercise compared outturns with projections at the start of the program for a three-year horizon (*T*, *T*+1, *T*+2). Because of data gaps in MONA, we examined projected and actual U.S. dollar values of aid flows in the fiscal accounts of staff reports for completed ESAF/PRGF arrangements in 20 sub-Saharan African countries.

The following are the main results (Appendix 3):

- Aid flows were projected to decline ("taper out") between the first and third year of the program in about three-fourths of cases. In half the cases, the magnitude of the projected decline was less than 1 percent of GDP, but in 10 percent of the cases projected declines exceeded 2 percent of GDP.
- For the first year of the program the direction of differences between projections and actuals are equally divided: in half the cases projections exceeded actuals and in the other half actual aid exceeded that projected. In most cases, the differences were less than 1 percent of GDP.
- Using the 20 case studies in sub-Saharan Africa, we find actual disbursements exceeding projections by more than 20 percent in a relatively small number of cases—between 2 to 5 cases depending upon the time horizon chosen. In fact, we observe a higher number of cases where projections exceeded actual disbursements by more than 20 percent (6 to 9 cases, depending on the time horizon chosen).

In summary, the data show that program projections of aid do tend to decline over the medium term in a majority of cases, albeit generally at a modest pace. However, on average, this does not appear to constrain aid flows on a year-to-year basis in programs that remain broadly on track. None of the evidence quoted here suggests that arrangements systematically underestimate aid flows in the outer years in program projections. However, the relatively simple analysis used here cannot answer the question—which goes beyond the scope of the current evaluation—whether more ambitious public spending (and deficit) targets, linked to poverty reduction, could have resulted in the mobilization of additional concessional external financing.

[5]We looked at program years for which MONA had data on both projections and outturns—mainly arrangements that remained on track over successive years. This reduced the sample size to 40 observations The outturn data for a particular program year was obtained from data reported in connection with a subsequent arrangement. Cases where there was a break in the series of one year or more between successive arrangements were dropped from the sample. Thus the sample was biased in favor of programs that remained broadly on track.

Social Spending Under IMF-Supported Programs: Cross-Country Evidence

Past IMF staff studies have investigated trends in health and education spending in developing countries. Gupta, Clements, and Tiongson (1998), using a sample of 118 developing and transition countries, find that since the mid-1980s real per capita spending on education and health has increased, on average, in developing countries but decreased in the transition economies. They observe that comparable increases can be observed for countries that had IMF-supported adjustment programs during the same period despite the fiscal consolidation often required by those programs.

In this section, we address the following question: What is the impact of the presence of an IMF-supported program on the level of social spending (other factors being held constant) relative to a situation without a program? For this purpose, we have investigated what happens to public sector social spending under IMF-supported programs using a broad sample of 146 countries in the 1985–2000 period.[6] Four different indicators were used for each type of spending: as a share of GDP, as a share of total government spending, as an index of real spending at domestic prices, and in U.S. dollars per capita.[7]

The basic statistical framework relates social spending in a particular country and year to the presence of an IMF-supported program that year and to a set of (control) variables that may also influence the level of social spending. The detailed discussion of methodological issues and results is provided in Appendix 4. We present here some basic descriptive statistics and our main conclusions.

Table 6.1 summarizes the mean values and standard deviations of each indicator for health and education spending. The size of the standard deviation relative to the mean indicates that there is considerable variability in the level of public spending on health and education.

One approach to determine the impact of IMF-supported programs on social spending is to compare periods with and without a program in a given

[6]A discussion of methodological issues and a presentation of results is in Appendix 4. For a more comprehensive report on the analysis and methodological issues underlying these findings, see Martin and Segura-Ubiergo (forthcoming). Social spending is measured on the basis of annual data on government spending on health and education using a database created by the Fiscal Affairs Department (FAD), and checked for accuracy by IMF staff from each country desk. See Baqir (2002) for a description and coverage.

[7]In the absence of a sector-specific price index, social spending was deflated by the general consumer price index. Expenditures in U.S. dollars were calculated at the annual average exchange rate, and deflated by the U.S. wholesale price index.

Table 6.1. Public Sector Social Spending Indicators

Indicator	Observations	Mean	Std. Dev.
Health spending			
As percent of GDP	1,452	2.2	1.5
As percent of total public spending	1,462	7.3	3.8
Per capita, at real domestic prices (index, country average 1985–2000 = 100)	1,418	100.0	30.0
Per capita, in U.S. dollars	1,424	6.1	9.4
Education spending			
As percent of GDP	1,452	4.2	2.0
As percent of total public spending	1,465	14.3	5.2
Per capita, at real domestic prices (index, country average 1985–2000 = 100)	1,413	100.0	25.3
Per capita, in U.S. dollars	1,419	10.2	14.8

Source: IMF, Fiscal Affairs Department.

country. This is reported in Table 6.2. In the large majority of countries for which data are available, there is no statistically significant difference in social spending between these two periods.[8] In the cases where the results are significant, the outcome depends on how spending indicators are measured. When spending in health and education is measured as a share of GDP or total public spending, we find there are more countries which show a significantly higher mean during program years than those that show a lower mean. However, the reverse is true when this spending is measured in per capita terms.

This type of comparison suffers from the obvious limitation that it attributes all the difference in program years to the fact of having a program. This is not a suitable counterfactual since there are other variables at work that affect social spending and their effect must be netted out.

To isolate the impact of an IMF-supported program on social spending, using the pooled cross-section time series data, we need a methodology that:

- Includes variables that have a direct effect on social spending, such as GDP per capita and share of school-age population. Not doing so would attribute to the presence of the IMF effects that are the result of these other variables (it is necessary to avoid a "*missing variable bias*").
- Recognizes that years with an IMF-supported program are not "normal" years, and that the special factors explaining the presence of a program could also, in principle, have an independent impact on social spending. For example, a country could seek an IMF-supported program as a result of an external shock (such as a sharp deterioration in the terms of trade) that may require a reduction in government spending with or without the presence of the Fund (i.e., it is important to take into account the *endogeneity* of IMF-supported programs).
- Takes into account that social spending tends to change sluggishly and is heavily affected by levels of spending in previous periods. This reflects not only that most programs are conceived as permanent or at least spanning several years, but also the political economy of budget allocation—most programs have constituencies that resist change. For these reasons, explanatory variables, including the presence of an IMF-supported program, are likely to have effects that are not instantaneous and may extend beyond one period (i.e., it is necessary to take into account possible problems of *serial correlation* and nonstationarity in the data series).

These problems have been addressed by using regression analysis in which we combine a series of explanatory variables that are directly expected to have an impact on social spending with the use of instrumental variables to model the presence of an IMF-supported program. (The estimated equations are reported in Appendix 4.)

The empirical results show that, on average, the presence of an IMF-supported program does not reduce social spending. In fact, the result shows that the presence of a program is associated with increased public spending in health and education measured as either a share of GDP, total spending, or in real terms compared with a situation without a program. However, the positive effects attributable to the program are short-lived. For these effects to be durable, they would have to be followed by further

[8] At least at the 90 percent confidence level.

Table 6.2. Number of Countries With and Without Statistically Significant Results

	Percent of GDP	Percent Total Spending	U.S. Dollars Per Capita	Domestic Real Prices Per Capita
Health spending				
Number of countries with (statistically significant) higher spending when there is an IMF-supported program	8	13	3	10
Number of countries with no significant difference between years with and without IMF-supported programs	78	76	83	75
Number of countries with (statistically significant) lower spending when there is an IMF-supported program	7	4	6	7
Education spending				
Number of countries with (statistically significant) higher spending when there is an IMF-supported program	7	11	1	8
Number of countries with no significant difference between years with and without IMF-supported programs	83	76	86	71
Number of countries with (statistically significant) lower spending when there is an IMF-supported program	5	8	6	14

Source: IEO staff calculations.

policy actions in these sectors beyond the program period. The results do not show any marked difference in the impact of programs supported by concessional or nonconcessional resources.

Figure 6.1 shows the estimated impact of a two-year IMF-supported program on education and health spending, using the regression results reported in Appendix 4, Table A4.1. The vertical axis provides point-estimates of the effect of a program relative to a situation without a program, all other factors being the same; the horizontal axis represents the timeline. Public spending in each of the health and education sectors increased by about 0.3 to 0.4 percentage point of GDP compared with a situation without a program. There is still a residual effect in the third year (when there is no longer a program), but this declines geometrically thereafter.

Whether this increase in spending sufficiently protects the most vulnerable groups during the program years will depend greatly on how well that increase in spending is targeted. If it is distributed according to past allocations—usually a high share spent in curative health or higher education and a high wage bill relative to recurrent inputs—the impact may be limited. If, on the other hand, it is used to fund targeted programs (old ones or new ones that can be activated during crisis) or to protect critical nonwage inputs (school supplies, school feeding programs, vaccines, and other critical medical inputs in basic health care), the impact could be much higher.

Role of the IMF in Connection with Social Expenditure and Social Protection

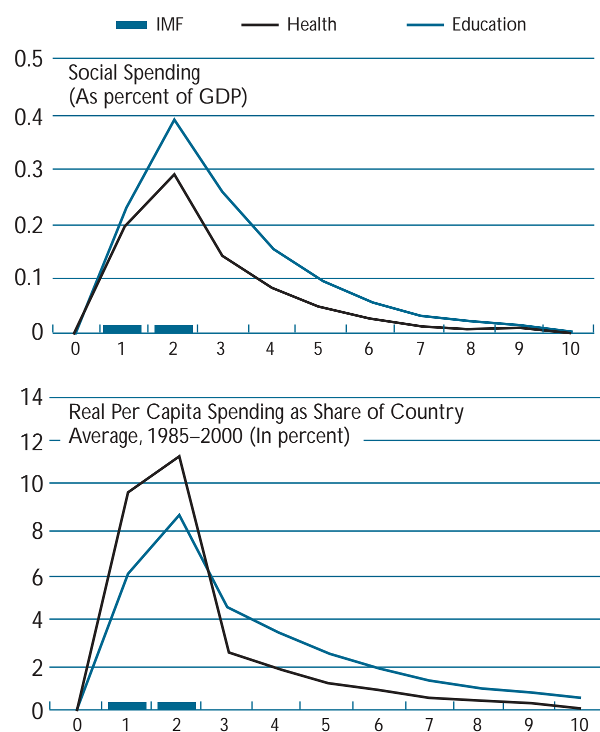
The role of the IMF vis-à-vis social spending has evolved as a result of a number of guidelines issued at different times. In 1991 the Managing Director issued guidelines to IMF staff directing that they should be explicitly concerned with the effects of economic policies on the poor and should discuss these concerns with government officials.[9] In 1997, new guidelines on social spending were issued to staff.[10] The guidelines emphasized the need for monitoring trends in this area and incorporating realistic targets into government budgets in the Letters of Intent on the basis of sector work by the World Bank (Box 6.1). In subsequent years, IMF management emphasized the need for a social pillar in the reform of the international financial architecture.[11]

[9]"Revised Guidelines on Poverty-Related Work," Office Memorandum from the Managing Director to Heads of Departments, March 8, 1991.

[10]"Guidelines on Social Expenditure," Office Memorandum from the Managing Director to Heads of Departments, May 28, 1997.

[11]Remarks by the Managing Director to UN ECOSOC Ambassadors, New York, June 31, 2000.

Figure 6.1. Estimated Impact of a Two-Year IMF-Supported Program



Source: IEO staff estimates based on regression coefficients.

In 1999, the Board discussed a paper on social issues in IMF-supported programs[12] in which the staff made proposals to (1) establish quantitative targets for education and health care spending and to strengthen efforts to monitor such spending; (2) occasionally set performance criteria on minimum spending thresholds; and (3) in some circumstances, monitor budget allocations for selected key inputs such as books and medicines. The Board discussion revealed divergent views on the subject. Several Directors urged caution, warning that the IMF should not allow its primary mandate to be diluted and pointed out that the IMF does not have the expertise needed to assess the quality of social spending and related issues and could best contribute to poverty reduction through its support of economic policies that provide a conducive environment for sustained growth. Some Directors felt that staff should assess, in the course of surveillance, the adequacy of social policy instruments, the performance of social safety nets, and the potential social ramifications of macroeconomic and financial policies, but others worried that this might detract from standard Article IV surveillance. Some Directors stressed the importance of efficient and well-targeted spending to ensure that gains in social indicators were commensurate with spending increases.

On the issue of incorporating social expenditures in program design, Directors considered that where social spending was critically low, structural benchmarks should continue to be used selectively to protect social spending and promote institutional reforms. However, while many Directors thought that such structural benchmarks should only be used in programs supported by concessional financing, others saw merit in also applying performance criteria to a broader range of IMF-supported programs. In establishing structural benchmarks, IMF staff would rely on input from the World Bank and other institutions to ensure that the targeting and quality of spending would remain optimal.

While the need for World Bank and IMF collaboration on social spending has been stressed on several occasions, it presents several operational problems in practice. These surfaced in the recent discussion by Executive Directors of proposals from the staff on collaboration with the World Bank on public expenditure issues.[13] Directors stressed that the IMF and the Bank should maintain a clear division of labor between the two institutions with the IMF taking the lead on the aggregate aspects of macroeconomic policy and their related instruments, and the Bank on issues relating to public expenditure composition and efficiency. They highlighted the need to better plan missions so as to reduce the burden on country authorities, better coordinate the different time frames of Fund and Bank work on public expenditure issues, and strengthen the collaboration with donors on country-led reform strategies. Directors also endorsed a framework that focuses on the articulation by the government of public expenditure reform strategies; an integrated and well-sequenced program of technical and financial assistance from development partners (including diagnostic work) to support countries' public expenditure reform strategies; and periodic reporting by countries of their performance in public expenditure policy, financial management, and procurement.

[12]Gupta, Dicks-Mireaux, Khemani, McDonald, and Verhoeven (2000) update the work presented to the Board in "Review of Social Issues and Policies in IMF-Supported Programs," EBS/99/171, August 27, 1999. The discussion in the next two paragraphs draws upon the summing up of the Board discussion.

[13]"Bank/Fund Collaboration on Public Expenditure Issues," SM/03/73, February 19, 2003. This paper does not explicitly address collaboration on social spending but the discussion is highly relevant.

Box 6.1. The 1997 Guidelines on Social Expenditure

The guidelines call for the following:

- IMF staff should use available fiscal data to keep track of main trends and developments in health and education spending and report these as memorandum items in fiscal tables in staff reports. Discussions on trends in social spending could be included in Recent Economic Development reports.
- IMF staff should rely on the sector expertise of other institutions in health and education and should, in particular, strengthen collaboration with World Bank staff. In those countries where health and education spending data are already available and relevant analyses from other institutions, in particular the World Bank, already exist, IMF staff should attempt to draw conclusions (on the basis of trends in the subject country and comparisons with other countries) regarding the level and efficiency of spending in health and education.
- IMF staff should rely on recent sector work by the Bank to incorporate realistic targets into government budgets and IMF-supported programs. These targets would not be expected to be performance criteria. It may be appropriate to encourage the authorities to incorporate such targets for health and education spending in the Letters of Intent for IMF-supported programs when the staff has examined the underlying analyses, and the targets are consistent with the overall macroeconomic framework and are monitorable.

IMF staff should continue to monitor developments in basic social indicators, such as poverty rates, infant mortality, life expectancy, illiteracy, school enrollment, and access to basic social services that are compiled by the World Bank and available online. In countries where such indicators are worsening or failing to improve in line with other developing countries, IMF staff should seek World Bank advice, and, if necessary, raise this issue with the authorities.

More recently, the emphasis on streamlining conditionality has raised new questions. Discussions with a number of staff suggest that there is uncertainty regarding how to interpret the 1997 Guidelines on Social Expenditure in light of the streamlining initiative.

In PRGF-supported programs, closer World Bank–IMF collaboration is mandated through the PRSP process, which calls for the monitoring of social and other poverty-reducing expenditures and for an explicit social impact analysis of major proposed policy reforms. Hence, in these countries, a framework for a more coordinated approach to social issues exists. However, for non-PRGF countries, there is a lack of clarity on how social policies should be handled. There is no PRSP-type framework and the World Bank may not have been involved in the social sector with the depth needed to deliver the relevant inputs on the short-term time schedule relevant for IMF operations. In these circumstances, the treatment of social issues in non-PRGF programs may well depend significantly on the emphasis provided by individual staff, the way they interpret the streamlining mandate, and the degree to which they collaborate with the Bank, itself dependent on the extent of readily available analysis done by the Bank. To assess what happens in practice, we examined a number of programs in depth.

A review of social issues in program design in 15 arrangements

The sample of 15 IMF-supported programs provides a basis for assessing how social issues are treated within the context of program design.[14] We posed a number of questions listed in Table 6.3, which also summarizes the results (elaborated in Appendix 5). Social spending issues are mentioned in almost all programs and changes in spending are noted in two-thirds of programs. However, little effort is made to sharpen the definition of social spending or to analyze the reasons behind trends. Only half the program requests that note changes in social spending actually analyze these changes. Few programs (other than in the PRSP/PRGF countries) establish explicit monitoring and feedback systems. Thus the empirical basis for identifying policy actions is often absent.

One difficulty is that social spending is not explicitly defined. Tables or boxes dealing with social spending in program documents typically associate social spending with education and health and sometimes tables indicate a single line titled "social spending" with no definition of the components.

About one-third of programs explore how to protect social spending, although typically at a very aggregate level of appropriations such as education spending. About 40 percent of programs used some conditionality in the form of benchmarks or indicative targets—none use performance criteria.

[14]One of these programs (Tanzania) was supported by concessional IMF resources and two (Senegal and Pakistan) by a mix of concessional and regular IMF resources. All the rest involved the use of IMF general resources only.

Table 6.3. Effectiveness in Identifying and Monitoring Social Spending in the Program Requests of 15 Selected Arrangements
(Percentage of cases where the answer to question is "yes")

Efforts at improving the empirical basis for policy	
Is social expenditure referenced at all?	93
Are changes in social spending noted?	67
Do programs include time series data on social spending?	67
Do programs define social spending clearly?	0
Are changes in social spending analyzed?	33
Efforts at identifying policies and actions	
Are there specific problems or issues identified?	80
Are there efforts to identify how social spending could be protected?	33
Are there any performance criteria or benchmarks in connection with social spending?	40
Did reviews follow up on issues raised in the program request?	100

Sources: IMF Staff Reports and IEO staff estimates.

Program reviews performed very well in following up whatever social issues were originally raised in the program request, and in many cases discussion of these issues was more extensive in the reviews than in the initial program request. For example, in Costa Rica the program request only briefly mentioned social issues and broadly discussed the need to strengthen the social safety net. The reviews, however, were more detailed and included more specific suggestions to achieve better targeting of social spending such as restructuring several agencies, decentralization, and encouraging the use of private suppliers of social services.

Similar patterns are found when examining comments from PDR and FAD during the internal review process. These comments often give feedback in this area, providing specific suggestions for the design and the support of priority social programs to protect vulnerable groups. However, most of these comments are concentrated in the reviews during program implementation and are, therefore, too late to influence the program design.

These results also suggest reasons why, despite good intentions, programs often fail to protect critical social spending. Programs recognize the need for action in the social sector but are vague about the specific types of spending that require protection. For example, in the case of the Philippines program, the staff report stated that "the staff urged the authorities to protect programs directed at poverty reduction in implementing the cuts. The authorities agreed, and explained that individual agencies had been instructed to reduce certain nonessential outlays (such as travel and training) by 50 percent. Agencies' revised spending plans are being reviewed with a view to protecting social programs as much as possible, especially those directed at poverty alleviation. Social programs would also be the first ones to be restored if fiscal developments during the year permit." Despite these good intentions, the proportion of the population served by various health programs declined, reflecting the absence of clear definitions regarding the specific critical programs to be protected, compounded by a lack of monitoring.

This picture, however, is not uniformly negative. The Algeria program, for example, defined very specific measures to revamp the social safety net in order to protect better the most vulnerable segments of the population via improved targeting. The program built on recommendations from an FAD technical assistance mission to introduce a public works program that would be self-targeting with a much lower remuneration than the minimum wage. Short-term unemployment would be dealt with by introducing an unemployment insurance mechanism to replace a system that imposed large severance payments on enterprises. Moreover, the authorities agreed to merge three other cash transfer schemes.

The use of conditionality to achieve social sector objectives was limited. Of the 15 programs examined, only 6 contain explicit social sector conditionality in the form of structural benchmarks and the implementation results were mixed. In the Algeria program, a structural benchmark was introduced to reform the social safety net through the introduction of a public works scheme and the benchmark was eventually met. In the Bulgaria program, a structural benchmark was set on improving the cost effectiveness of health care, and that benchmark was subsequently only partially met. For the Pakistan program, an indicative target was put on social and poverty-related spending, but the target was not met. The Senegal program included a structural performance criterion relating to budgetary allocations for the health and education sectors. A closer look at the criterion, however, reveals that it actually only called for an action plan and communication to IMF staff on the issue. In the Ukraine program, a benchmark was set

Table 6.4. The Ecuador Program: Imbalance Between Efficiency and Equity Measures Underpinned by Conditionality

Measures in the 2000 Memorandum of Understanding	Included as a Performance Criterion (PC)/Benchmark (B)?
Adjustments of prices	
Fuels	Yes (PC)
Cooking gas	Yes (PC)
Electricity rates	No
Other fiscal measures	
Eliminate temporary tariff surcharge	Yes (B)
Control over expenditure, including wage bill	Yes (PC)
Payment of domestic arrears	Yes (PC)
Tax measures	
Raise VAT and increase tax base	Yes (B)
Lower income tax threshold	Yes (B)
Reduce evasion	No
Reduce loopholes	No
Improve tax administration	No
Reduce earmarking	Yes (B)
Elimination of nuisance taxes	Yes (B)
Consumption tax on gasoline	Yes (B)
Social measures	
Adjustment of *Bono Solidario*	No
Improve targeting of *Bono Solidario*	No
Nutrition and family programs	No
Community programs	No
Education programs	No
Increase social spending if revenues allow	No

Source: Ecuador program documents.

on specific reforms in the health and education ministries and that benchmark was also met, although some slippage occurred after the benchmark was removed from the program. The Venezuela program had structural benchmarks calling for legislation to reform the severance payment system and strengthen the social safety net. These were implemented but with delay.

There are situations where poorer groups have not only been adversely affected by output declines and devaluations in crisis periods prior to programs, but also by fiscal and price adjustment measures included in programs for macroeconomic reasons but which may have second-round adverse effects. The Ecuador program was well aware of this phenomenon and it supported the government's plan to index the preexisting cash transfer program (*Bono Solidario*) and other poverty programs to offset negative effects on the poor. However, although there was clear conditionality on the pricing of fuels, spending control, and raising the VAT, none of the social measures in the Letter of Intent with the purpose to offset these effects was incorporated as a structural benchmark (see Table 6.4).

A critical issue for program design is whether critical programs can be protected at affordable cost and in a manner which can be effectively monitored. This is certainly possible but it requires a high level of control over institutional management to implement these measures of protection. Box 6.2 shows how public hospitals in Ecuador adjusted to the 1998–99 crisis prior to the program. The wage bill and personnel expenses were protected but free provision of drugs to patients and even food for inpatients declined sharply relative to spending on personnel. Nonwage inputs—which are a small share to begin with (only 20 percent of hospital spending)—were squeezed. In principle, it should be possible to protect these items without jeopardizing any macroeconomic target in any standard program. However, doing so requires identification of critical programs and spending categories prior to the crisis and the ability to ensure that the relevant allocations are effectively protected when they come under pressure in crisis situations.

There are examples of cost-effective and targeted programs that could be protected at low fiscal cost in case of a crisis. One example comes from Tanzania (see Box 6.3), where well-targeted health intervention with an emphasis on children was implemented in a pilot program covering two districts at a cost of less than $2 per capita. Another example is the Pro-

Box 6.2. How Public Hospitals in Ecuador Adjusted in a Time of Crisis

As a result of a series of external shocks and a domestic banking crisis, Ecuador experienced a macroeconomic crisis of major proportion in 1999. Output declined by 7.5 percent, inflation accelerated to approximately 60 percent a year, and the sucre/dollar exchange rate almost doubled.

While nominal public sector wages increased by 34 percent between 1998 and 1999, the health budget only increased by about 12 percent. Under these circumstances, how did a typical public hospital adjust when salaries accounted for about 80 percent of its operations and the cost of nonwage medical inputs went up with the devaluation? To answer this question, a sample of six large public hospitals in Quito and Guayaquil were visited to assess how they coped with the crisis. They accounted for about 12 percent of the total number of hospital beds nationwide.

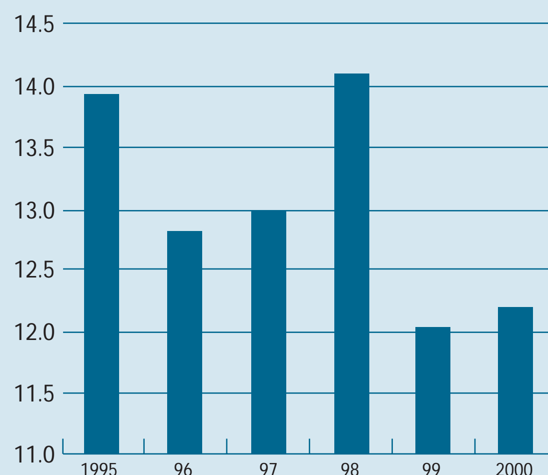
The major finding was that the sharp erosion in real budgets in 1999 translated into a reduction of nonwage medical inputs and maintenance of equipment. Consequently, hospitals were forced to cut back care to patients. In three of the four hospitals that provided data, outpatient services declined 26 percent to 37 percent.

In addition, the number of drug prescriptions dispensed declined very sharply in three hospitals, by amounts ranging from one-half to four-fifths, and increased by about 10 percent in those hospitals where some cost recovery was feasible. Independent data for the overall public health system show a decline of about 14 percent in the total number of prescriptions dispensed by the entire system (see figure).

For some of the hospitals visited, data were obtained on the number of food rations received by the hospital staff versus patients. In the Quito hospital, rations for patients were reduced during the crisis—sometimes severely—while those for staff remained relatively constant. Only in one Guayaquil hospital were food rations maintained thanks to additional funding received by the hospital to mitigate the impact of El Niño on the coastal areas.

This example illustrates that the protection of small but critical nonwage budgetary items under fiscal adjustment is a major challenge in the design and monitoring of adjustment programs.

Number of Prescriptions Dispensed
(In millions)



gresa Program in Mexico. Poor rural families received cash transfers, school supplies, and nutrition supplements conditional on children's school attendance and regular preventive health care. The program has reached about 2.5 million households at a cost of about 0.2 percent of GDP. Budgetary shocks that threaten these allocations can be protected at low cost and with little impact on the overall fiscal program. In summary, if countries introduce beforehand well-targeted social programs, they can easily be protected or activated at low fiscal cost in a crisis situation.

The experience of Chile (not part of our evaluation) is of general interest for middle-income countries. Not only has Chile been effective in protecting critical programs such as children's basic health care and nutrition, but it has also been able significantly to realign the budget toward social spending while improving the incidence of public spending towards the lower-income population. This has been accomplished without unduly increasing the tax burden. That tax burden is about 19 percent of GDP, a product of moderate tax rates and good collection. About 70 percent of spending in basic social services and cash assistance is focused on the first two quintiles of the population. These achievements have been the product of many years of institutional reforms and political consensus regarding these policy priorities, and it provides a good reference point of what is possible.

In addition to examining social sector issues in the 15 main programs chosen for this study, we went a step further in order to evaluate the latest arrangement for 8 of the 15 countries for which there was a more recent program (these include the Algeria SBA 1995, Bulgaria SBA 2002, Jordan SBA 2002, Pakistan PRGF 2001, Peru SBA 2002, Romania SBA 2001, Tanzania PRGF 2000, and Uruguay SBA 2002 programs). We adopted identical criteria to those

Box 6.3. Protecting Critical Programs Is Not Costly When Programs Are Well Targeted

An experimental health intervention in Tanzania shows that small additional resources devoted to health care in a poor country can alleviate the burden of disease if carefully allocated. The intervention was carried out in two rural districts by the Tanzanian Essential Health Intervention Project (TEHIP), a joint venture of Tanzania's Health Ministry and Canada's International Development Research Centre (IDRC).[1]

The key innovation was to focus financial resources on diseases that imposed the highest burden on the population. It was found, for example, that a cluster of childhood problems such as malaria, pneumonia, diarrhea, malnutrition, and measles accounted for 28 percent of disease in the districts, but only received 13 percent of the local health care budgets. An additional $2 a head allocated to the district's health care budget was to be spent on diseases with the largest social cost based on years of life lost. The results thus far have been dramatic. Infant mortality fell by 28 percent from 1999 to 2000. The number of deaths prior to five years of age dropped by 14 percent. There is no evidence of similar improvements in that period in nearby districts or in Tanzania overall.

[1]Reported in *The Economist*, August 17, 2002.

These are the types of programs that need to be protected under macroeconomic shocks that put pressure on public finances. It is clear that IMF-supported programs could make room for such interventions. However, making sure public expenditure management systems are able to deliver resources to desired destinations depends on local knowledge and will require support from the World Bank. It is not possible to set up such monitoring and delivery systems within the short time frame in which the negotiation and implementation of an IMF-supported program takes place. Nor is this an area where the IMF has the necessary expertise.

To deal with such problems of a potential mismatch of time frames, the IMF needs to encourage the authorities, independently of the negotiation of a particular IMF-supported program (and probably with support from the World Bank and other external partners), to (1) identify core budgets that would be protected in case of budget cuts; (2) develop public expenditure management systems capable of monitoring the flow of resources to critical programs in real time; and (3) protect the cash flow to items in the core budget during times of fiscal pressures. In countries like Tanzania, the framework of the PRSP exists to address such issues, but the approach to be taken is less obvious in non-PRSP/PRGF cases.

used to assess the treatment of social issues in the original 15 IMF-supported programs. Results show that the more recent programs exhibit slight improvements in categories such as noting and analyzing changes in social spending, identifying specific social spending issues, and actions to protect social spending. In 3 of the 8 programs, structural benchmarks were used to support social protection measures. At the same time, there is little change or even a slight deterioration in presenting a series of social spending data. This suggests there is still room for considerable improvement.

Conclusions

It is clear from our evaluation that protection of social spending on critical and well-targeted programs in the social sector can play an important role in protecting vulnerable groups from adverse shocks and budgetary retrenchments at fairly low cost. This emphasis is also consistent with the IMF Articles of Agreement (especially Article I (v)) and with commitments made in the follow-up of the 1995 World Summit for Social Development (see IMF, 2000a). Efforts should, therefore, be made to build such elements into program design wherever possible. However, a framework is necessary that takes account of four operational constraints. (1) To be effective, and acceptable, policies in this area must be truly homegrown and fully owned domestically; the initiatives must, therefore, come from the country. (2) Since the IMF does not have expertise on social sector issues, nor is this an area of its comparative advantage, inputs from other agencies, especially the World Bank (and possibly also others), are critical. (3) There is a mismatch of time frames between the short-term nature of IMF programs and the longer-term time frame needed for building institutions and budgetary systems that can provide social support in times of crisis effectively. (4) Finally, it is necessary to ensure that incorporation of social protection system does not contradict the recent streamlining initiative by leading to an overload of conditionality.

In the case of low-income countries, the PRSP framework could potentially meet these requirements. The extent to which this is actually achieved will be separately examined in the ongoing IEO evaluation of the PRSP/PRGF experience. However, there is at present no framework for non-PRGF eligible, predominantly middle-income countries that would ensure identification of critical and homegrown social sector support programs that could be used as mechanisms for social protection at the time of crisis.

The PRSP framework is obviously not appropriate for middle-income countries, but in the absence of any framework there will be a growing divergence between the way social issues are treated between PRGF and non-PRGF countries. It is, therefore, necessary to revisit the 1997 guidelines with special reference to what IMF staff should do consistent with the overall operational constraints listed above.

Some elements of a workable approach can be readily identified. First, the mismatch of time frames suggests that work in this area must be undertaken not at the time of crisis but much earlier as part of normal surveillance. In order to encourage a homegrown initiative, the IMF could request governments to consider identifying critical social sector programs that could serve as effective social safety nets that could be intensified in the event of crisis. The IMF could encourage countries to approach the World Bank for assistance in this area. The IMF on its part, consistent with its mandate, could report on the authorities' responses in this area and monitor programs in developing social safety nets.

Building on recent initiatives (such as the call for increased coordination on public expenditure management (PEM) issues), both institutions could agree with the authorities on the reforms that would need to be tackled and an appropriate sequencing. Where joint efforts are required, for example in public expenditure management, a work program in these areas would be jointly established. On the basis of the resulting joint effort, the IMF and the World Bank would assist the authorities in setting up mechanisms to track critical social spending throughout the budget and identify ultimate allocations including to local governments where a significant amount of spending is decentralized. In this regard, establishment of better and more transparent monitoring systems is probably one of the major contributions that can be made to encourage homegrown policy initiatives in this area.

Chapter 7

Fiscal Reforms in IMF-Supported Programs

Fiscal adjustment in IMF-supported programs typically includes an agenda of fiscal reforms, and in this chapter we focus on the experience with such reforms on the basis of the sample of 15 programs. We then turn to the process of learning from the past and the role of surveillance in monitoring reform and its link with program design.

Fiscal Reforms in Programs: An Overview

Each program typically includes a number of reform measures in the fiscal area. The 15 programs studied for this evaluation identified 153 specific fiscal-related reform measures, of which 101 were subject to conditionality (divided into 79 structural benchmarks and 22 performance criteria).[1] In this chapter we present an overview of these measures in order to identify the relative emphasis placed on different reforms. We also provide an assessment of the success in implementation in different areas based on the assessments reported by staff in program documents.

The universe of reforms can be divided into nine categories: tax policy; tax administration; wage bill and civil service reforms; social sector spending; other spending issues; public enterprise reform, privatization, and private sector development; social security and pensions; organizational reform; and pricing policy of public utilities. Box 7.1 describes the typical reform measures in each category.[2]

[1]Some qualifications regarding the universe of reforms are necessary. First, for programs possessing extensive reform agendas, such as the Bulgaria and Ukraine programs, we narrowed down the number of reform measures to a subset of reforms representing the major areas of emphasis. Second, for programs possessing obvious groupings of intricate and interrelated reform measures, we collapsed various measures into one all-encompassing measure; for example, the Uruguay program requires various measures relating to the publishing of fiscal data, other reports, and studies. These have been consolidated into an umbrella "transparency and disclosure" reform measure.

[2]Due to the small sample size of the measures supported by performance criteria, we collapse the structural benchmarks and performance criteria into one single group referred to as measures supported by "conditionality."

Table 7.1 presents in summary form the frequency of occurrence of the different types of reform measures in the 15 programs as well as the frequency of occurrence of those supported by conditionality. The areas supported by conditionality follow a pattern similar to the overall universe of measures. Tax policy, public enterprise reform, and privatization are the areas of largest emphasis for conditionality, followed by organizational reform, wage bill and civil service reform, and tax administration. Social sector and other spending reforms are typically little emphasized in conditionality.

The data also suggest that programs tend to emphasize revenue-related reforms over those related to spending, with a focus on tax policy relative to tax administration. Quasi-fiscal issues, particularly public enterprises, receive more coverage than some core fiscal issues, such as spending reform. The emphasis on quasi-fiscal issues may be attributed to efforts aimed at redefining the overall role of government, which is particularly evident in the sample of transition economies. It may also reflect that earlier adjustment efforts focused on bringing extrabudgetary activity into the central budget (e.g., extrabudgetary funds, public enterprises, and implicit and explicit guarantees in lieu of explicit subsidies). The stress on revenue and the limited attention paid to reallocating or reforming nonsocial spending may also be the result of the short horizon of programs and concentration of IMF expertise.

A number of programs incorporated measures that reduced the short-term deficit but did not reduce fiscal vulnerabilities or improve sustainability. Examples include across-the-board cuts that usually spare the wage bill (e.g., the Philippines) and increasing tax rates on a narrow base, such as raising already very high social security contributions (Romania).

Tax reform focuses much more on introducing or expanding VAT or increasing VAT rates as well as reducing trade tariffs, with relatively less attention paid to income and property taxes.[3] Less attention is also

[3]Property taxes are usually levied by local governments while the IMF focuses on the central government. However, to the extent that local governments receive significant transfers, the central government has leverage to press for a more aggressive use of property taxes.

Box 7.1. Public Finance Reform Areas[1]

The 153 fiscal reform measures reported in the staff reports for the 15 IMF-supported programs are divided into 9 reform areas:

Revenue

1. Tax policy. (i) Introduction of the VAT (Jordan and Tanzania) or modifications to the VAT such as widening the base (Algeria, Bulgaria, the Philippines, and Ukraine), or rate increases (Ecuador and Senegal); (ii) introduction or expansion of other consumption taxes including excises and taxation of petroleum products (Ecuador, Egypt, Pakistan, the Philippines, and Romania); (iii) reduction of taxes on international trade (Algeria, Bulgaria, Egypt, Jordan, Pakistan, Peru, the Philippines, and Ukraine); and (iv) income tax reform (Ecuador, Egypt, Jordan, Pakistan, the Philippines, and Ukraine).

2. Tax administration. (i) Measures aimed at large tax payers (Bulgaria, Peru, and the Philippines); (ii) improving identification of tax payers (Bulgaria and Pakistan); (iii) strengthening enforcement of collections (Bulgaria, Peru, Romania, and Tanzania); and (iv) personnel training (Jordan).

Expenditure

3. Wage bill and civil service reforms. (i) Wage bill controls (Algeria); (ii) limiting wage increases (Algeria, Peru, Romania, Tanzania, and Uruguay); (iii) limits or cuts in employment (Costa Rica, Egypt, the Philippines, Tanzania, and Ukraine); (iv) legislative action to change civil service statutes (Bulgaria, Costa Rica, and Venezuela); and (v) formulation of reform proposals (Pakistan and Ukraine).

4. Social sector spending. (i) Reform of social sector subsidies (Algeria); (ii) improved targeting (Algeria, Bulgaria, and Ukraine); (iii) improvement or introduction of social safety net (Algeria and Venezuela); and (iv) increase and/or rationalization of welfare spending (Bulgaria, Pakistan, Peru, and Ukraine).

5. Other spending issues. (i) Rationalizing public investment (Algeria); and (ii) reducing spending (Uruguay).

Quasi-fiscal

6. Public enterprise reform, privatization, and private sector development. (i) Restructuring public enterprises (Algeria, Jordan, Senegal, Uruguay, and Venezuela); (ii) privatization (Bulgaria, Egypt, Jordan, Pakistan, Peru, Romania, Senegal, Tanzania, and Ukraine); and (iii) encouraging private sector entry to areas dominated by the state (Costa Rica, Jordan, and the Philippines).

7. Social security and pensions. (i) Ensuring the viability of pension systems (Bulgaria, Peru, Senegal, Ukraine, Uruguay, and Venezuela).

8. Organizational reform

(i) Transparency in government accounts or budgeting (Bulgaria and Pakistan); (ii) improved coverage of budget including extrabudgetary funds (Bulgaria and Ukraine); (iii) reduced earmarking (Bulgaria); (iv) improved public expenditure management such as budgeting procedures (including multiyear), controls, and audit (Bulgaria, Jordan, Pakistan, Peru, the Philippines, Ukraine, and Senegal); and (v) creation or revamping of institutions including to manage debt or natural resource–related revenue (Venezuela).

9. Pricing policy

(i) Decontrol or raising of energy-related prices with a fiscal impact (Ecuador, Egypt, the Philippines, Senegal, and Venezuela).

[1]The examples in this box are illustrative and not meant to cover all 153 reform measures.

given to reducing tax exemptions and evasion of income taxes and customs duties. For example, in the Tanzania program, the reduction in import duties was not accompanied by equivalent efforts to reduce tax evasion in the ports.

The Ecuador program provides a dramatic example of what can be achieved when a determined effort is made to reduce evasion broadly, rather than relying on a VAT rate increase. This effort started prior to the Ecuador program and yielded significantly higher revenue than those envisaged in the program. In fact, these unprogrammed increases in revenue owing to improved tax collection were significantly higher than those expected from the programmed increase in VAT rates (Box 7.2).

Often, tax administration reform has focused on the technology side (information systems, manuals for training, and the like) rather than on politically demanding action, including steps within the purview of the executive branch or legislation that would empower tax administrators to collect tax arrears, forcefully pursue tax evasion, and be better protected from political influence. There appears to be significant scope for reorienting efforts to a more vigorous attack on tax evasion and exemptions parallel with efforts to increase VAT rates or broadening the base.

On the spending side, conditionality has concentrated on short-term quantitative targets to reduce public employment, or cap public sector wage increases, or across-the-board spending cuts. The benefits are usually short-lived because of the easily reversible nature of these measures compared with the reorientation of public spending and civil service re-

Table 7.1. Distribution of Areas of Fiscal Reforms and Those Supported by Conditionality
(In percent)

	Areas of Fiscal Reform	Measures Subject to Conditionality
Tax policy	26	25
Tax administration	14	9
Wage bill and civil service reforms	12	10
Social sector spending	7	5
Other spending issues	2	3
Public enterprise reform, privatization, and private sector development	19	25
Social security and pensions	4	4
Organizational reform	11	12
Pricing policy	5	7
Total	100	100

Source: IEO staff calculations, based on program documents.

form geared to improve efficiency and link pay to productivity. Except for PRGF-supported programs, there is relatively little emphasis on improving pro-poor public spending beyond vague statements concerning better targeting.

The internal review process usually addresses several of the areas of weakness identified earlier, such as the need to look also at income taxes, spending reallocations, and, perhaps most important, the need for determined actions by the executive in the areas of reducing tax exemptions, limiting tax incentives, and taking concrete actions against tax evasion and tax arrears. But again, these comments come mainly during the review of program implementation, rather than at an earlier stage when they would have more impact on program design.

In summary, the overall picture that emerges is one of heavy emphasis on the revenue side relative to spending reform. On the revenue side, the accent has been on increasing the yield from VAT/consumption taxes. This may reflect the need for measures that quickly yield revenue increases. However, other measures that could also provide important revenue in the short run, such as forceful efforts at collecting tax arrears and reducing tax evasion and exemptions, have received relatively less attention. On the spending side, most measures aim at capping the public sector wage bill through quantitative targets. Less emphasis has been given to reallocating public spending and launching durable civil service reforms. This emphasis may again reflect the mismatch between the quantitative targets and the short length of programs on the one hand, and the time required to complete institutionally and politically difficult reforms on the other hand. Many of these conclusions have also emerged from past staff assessments of cross-country experience in fiscal reform (Abed and others (1998), Mackenzie and others (1997), and Schadler and others (1995a and 1995b)).

Progress in Implementing Reforms

This section presents an assessment of the extent to which programs have been effective in implementing the reform agenda discussed in the previous section. For this purpose, we tracked each of the 153 reform measures described earlier.[4] Progress was classified into three categories: "significant," "partial," and "little" on the basis of staff's own evaluation of progress as reported to the Board in program review documents covering the life of the arrangement.[5] An index

[4]One caveat to our findings: This analysis has been handicapped by lack of consistency in following up and/or reporting progress with reform. This limitation has introduced an element of subjectivity in interpreting progress in implementing the agreed structural measures.

[5]"Significant progress" indicates that by the end of the program most of the agreed reform was enacted. For example, the first review of the Bulgaria program reports that the largest extrabudgetary funds were incorporated into the budget, as envisaged. "Partial progress" indicates that the agreed agenda remains to be implemented but there was noticeable movement in a positive direction. For example, the Pakistan program called for improved spending monitoring based on a variety of transparency, governance, and accounting measures. However at the end of the program, the review stated that "reconciliation of especially provincial spending remains too slow, resulting in large amounts of spending remaining unclassified for too long, thus hampering proper expenditure management and prioritization." On the other hand, the program resulted in improved fiscal transparency, such as publishing reconciled public accounts. "Little progress" suggests change that is barely perceptible, if at all. For example, the Egypt program envisaged phasing in the extension of the input crediting mechanism to capital goods under the general sales tax from January 1997. However, this reform was delayed more than once owing to lack of parliamentary approval.

Box 7.2. Ecuador: What a Determined Head of Tax Administration Can Do

The success of the Ecuador program (SBA 2000) in increasing revenue was due to both an increase in the VAT from 10 percent to 12 percent, but, more important, to radical improvements in tax administration that started in 1997.

In the 1993–97 period, tax revenue averaged only 6.7 percent of GDP, including 3.5 percent of GDP from VAT. A myriad of loopholes, exemptions, sophisticated evasion, and tax erosion prevailed in the tax system. The country did not have a modern functioning tax administration. The Tax Collection Department of the Ministry of Finance lacked a tax accounting system. It relied on outdated tax forms and did not have any information cross-checking system.

After 1998, tax collections increased dramatically—rising by 80 percent from 1998 to 2001, mainly because of sharp increases in value-added and income tax collections. The collection efficiency of the VAT increased from 42 percent to 68 percent. Of the 4 percentage points of GDP improvements in VAT collections, about one-fourth can be attributed to the increase in the VAT rate, with most of the increase reflecting improved tax administration.

What had changed so abruptly? In mid-1997, the Internal Revenue Service (SRI) was created as an autonomous government agency. The first year was dedicated to basic reforms to the old tax collection department inherited from the Ministry of Finance, but progress was limited. Following the nomination of a new head in September 1998, a massive process of reform started. An important share of the personnel was dismissed and new staff was hired, and incentives and compensation were improved due to the autonomous nature of the agency complemented by improvements in technology and training. The agency started a forceful process to control evasion, such as surprise visits to enterprises to check invoices and the vigorous implementation of penalties, including closures of enterprises. The overall process was supported by technical assistance from the Inter-American Development Bank.

The lesson is that institutional changes accompanied by determined enforcement can improve collection by amounts significantly higher than increases in tax rates. These changes take time and need to be encouraged by continuous efforts during noncrisis periods. However, legal changes per se do not suffice if, due to political interference, heads of administrations are inhibited from using the available legal tools.

Tax Collections

	Average 1993–97	1998	1999	2000	2001
			(In percent of GDP)		
Total tax collection	6.7	7.4	10.0	12.3	13.2
Income tax (personal plus corporate)	2.0	2.1	0.7	2.0	3.2
Value-added tax	3.5	4.2	4.5	6.8	8.2
Excise consumption taxes	0.7	0.6	0.6	0.7	1.0
Other[1]	0.5	0.6	4.3	2.0	0.7
			(In percent)		
Collection efficiency of VAT[2]	35	42	45	57	68
Memorandum item:					
Nominal VAT rate	10	10	10	12	12

Sources: SRI and Banco Central del Ecuador. In January 2003, a new official GDP series starting in 1993 was introduced, with upward adjustments in GDP figures by about 15 percent to 20 percent. Expressing tax collection as a share of this new GDP series reduces the level of tax collection as a share of GDP. However, it does not change the significant trend toward improvements in tax collection as a share of GDP.

[1]In 1999–2000 includes the 1 percent capital transactions tax. Between January and April 1999, income taxes were abolished and replaced by the 1 percent tax.

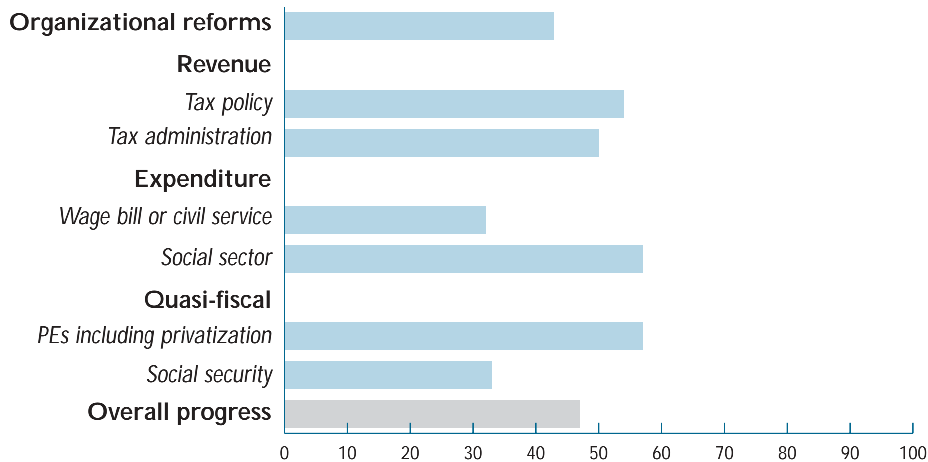
[2]The ratio of actual VAT collections over GDP times the legal rate.

was constructed to measure performance in implementation in each area by assigning a weight of zero to cases of "little progress," 0.5 to "partial progress," and 1.0 to "significant progress."[6]

Our assessment focuses on reform measures highlighted in IMF-supported programs as reflected in documentation presented to the Board. This only provides a partial view of the total efforts of the IMF in promoting reform of public finances and strengthening fiscal systems. A more complete picture would need to consider medium-term efforts of technical assistance (TA) to address key problems of fiscal systems. We have not done so here because this evaluation concentrates on fiscal adjustment under spe-

[6]The index can be interpreted in one of two ways. For example, an index value of 50 percent could indicate that on average about half of the reform measures were successfully implemented. Alternatively, it could also be interpreted as all reform measures showing only partial progress in implementation.

Figure 7.1. Index Indicating Implementation Progress in 153 Fiscal Reform Measures in 15 IMF-Supported Programs
(In percent)



Source: IEO staff calculations based on program documents.

cific IMF-supported programs. IMF technical assistance will be the subject of a separate forthcoming evaluation by the IEO.[7]

The results are summarized in Figure 7.1 for several reform areas (Appendix 6 provides the detailed results).[8] Overall, the index of implementation ranges between 30 percent and 60 percent—indicating a mixed picture and a sense of partial success at best. Tax policy and tax administration, social sector, and public enterprise reform seem to fare better. Social security and wage bill or civil service reform tend to perform worse.

Areas of strength and weakness

The index provides a measure of "average" performance. To get some idea of variation we have also looked at the distribution of programs according to progress achieved (Figure 7.2). The highest degree of success in terms of significant progress in implementing reform is achieved in the social sector area. Approximately 40 percent of reforms in this area were implemented with significant progress. However, even in this relatively successful area there was significant variability. An example of success includes the introduction of a public works scheme and unemployment compensation under the Algeria program. The Ukraine program illustrates partial progress owing to delays and incomplete implementation of plans to improve the efficiency of health and education spending and targeting of allowances. The finding that implementation of reform in this area was relatively successful is not inconsistent with the earlier findings (Chapter 6) that specific social sector reforms were addressed in less than half of the programs. This means that when these reforms were indeed addressed, their implementation was good in relation to other reform areas.

In the middle of the performance scale, we identified three areas in which about 30 percent of reform measures showed significant progress. These are organizational reform (including public expenditure management), public enterprise reform (including privatization), and tax administration. However, even across these areas, there was considerable variation.

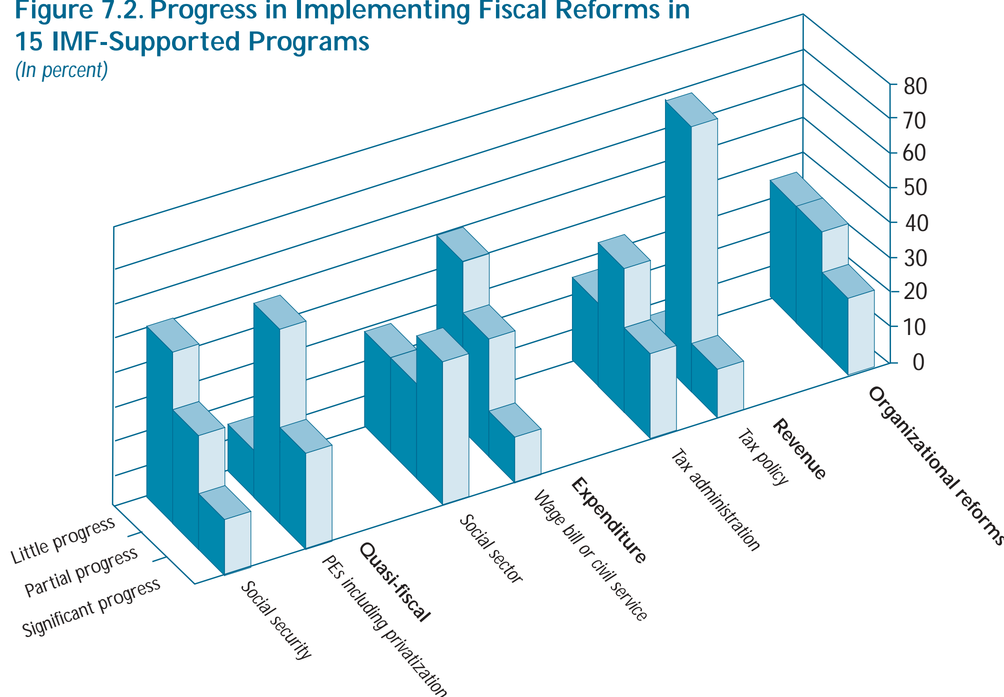
In organizational reform, the elimination of earmarking under the Bulgaria program is an example of success. Limited progress with organizational reform is illustrated by the failure to establish a debt redemption fund under the Venezuela program or to include extrabudgetary funds in the budget under the Ukraine program. We also consider progress in the Senegal program to be limited, despite having met a basic objective of the program (preparation of documents proposing how to improve public expenditure management).[9] Partial progress with organizational change under the Philippines program is reflected in

[7]See Appendix 7 for a summary of the IMF's fiscal TA in the 15 countries.

[8]The discussion excludes the "other spending" area given that it is only covered in three programs and the sample size is thus too small to draw general conclusions. Pricing policy is also excluded as it is not a core fiscal area.

[9]This is the one case where we are more critical than the assessment of the staff due to the undemanding measures required.

Figure 7.2. Progress in Implementing Fiscal Reforms in 15 IMF-Supported Programs
(In percent)



Source: IEO staff calculations based on program documents.

movement to a three-year budgeting framework but only limited results in reviewing the devolution of funds and responsibilities to local government units.

Within public enterprise reform, significant progress includes accelerating privatization under the Romania program. Partial progress includes the unfinished preparation of a plan to restructure public electricity utilities under the Venezuela program.

Regarding tax administration, examples of limited progress include lack of improvement and follow-up on measures to increase penalties for tax evasion and close loopholes in the Peru program; the failure to collect revenue in cash and abstain from netting out operations under the Ukraine program; limited progress in meeting structural benchmarks to strengthen tax administration and taxpayer registration under the Pakistan program.

The proportion of substantial implementation of structural reforms was lowest—under 20 percent—in areas such as civil service and/or wage bill reform and social security (including pensions), as well as tax policy. Successes include limits on the wage bill under the Tanzania program and pension reform under the Peru program to cover unfunded liabilities and issue pension bonds. Partial progress is exemplified by the Costa Rica program that met targets for reducing public sector employment but not those for approval of a Public Employment Law; and by the submission under the Uruguay program of a law to reform special pension funds for some, but not all, groups. On the wage bill, limited progress was achieved, for example, under the Egypt program, which failed to achieve the targeted 2 percent annual reduction in employment or under the Peru program, which failed to contain wage increases to an average 12 percent. An example of limited progress on social security reform comes from the Venezuela program, which failed to result in measures to improve the finances of the IVSS (Social Security Institute).

Why institutional reforms have often been so intractable in IMF-supported programs—examples from the case studies

The previous section showed that significant progress in fiscal reform areas has been limited—in no area did it exceed 40 percent of cases. Insufficient institutional reforms in areas such as tax administration, reallocation of spending, public expenditure management, and civil service reform results in insufficient progress in improving the long-term equity and efficiency of public finances and the flexibility of fiscal systems in response to shocks.

The case studies bring out some of the reasons progress in these areas has been limited. Often it is due to an excessive emphasis in meeting short-term quantitative targets rather than focusing on critical institutional changes that might extend beyond the

end of the program. It is largely the result of a mismatch of time frames, for example, the short horizon of programs relative to the time needed to complete these institutional reforms.

As pointed out by Tanzi (2000a), many developing countries now face important second-generation fiscal reforms that focus more on improving institutions than reforming policies. These institutional changes require significant time compared with first-generation policy reforms. The IMF needs to address the resulting mismatch of time frames since the benefits from the first-generation reforms need to be sustained through second-generation reforms. Such reforms may need to be broken down into several steps: some of them can be started at the outset of the program with enough determination from the executive branch; others will require time to the extent they call for legislation and improvements in the implementation capacity of agencies. We elaborate below with some specific examples.[10]

Examples on the revenue side

During the 1998 Philippines program, tax collection deteriorated owing to governance problems that remained unresolved for the duration of the program—reversing earlier painfully acquired progress.[11] Moreover, the inability to reduce tax evasion remains critical today, as noted by the December 2002 Post-Program Monitoring Mission.

Increases in tax rates of "easy to collect taxes" may not be effective when such rates are already high and the tax base is low. For example, in the Romania program, there were diminishing returns to raising already high social security taxes imposed on a low base (see Box 7.3). That lack of flexibility could have been prevented if long-term reforms to widen the tax base and reduce evasion had been pursued more forcefully over time.

There are also occasions when total revenue might fall if tax reform that (rightly) reduces trade taxes and excessively high statutory income and corporate tax rates is not accompanied by measures to improve collection and reduce exemptions. Reductions in tax rates are institutionally easy—they are stroke-of-the-pen reforms with few losers. In contrast, improving collection requires politically demanding decisions and the development of strong independent revenue-collection agencies. For example, during the implementation of the Tanzania program, tax evasion increased in the ports as trade expanded and important tax exemptions were granted to importers of petroleum (Box 7.4).

Many reforms to improve revenue performance (both quantitative and qualitative) require different time spans and are subject to different constraints such as: (1) lack of support of the executive to encourage tax agencies to collect tax arrears and improve collections from well-known sources of tax evasion owing to lack of political will; (2) lack of legislation to empower tax agencies which hinders effectiveness even though the executive is willing to support the actions of these agencies; (3) implementation capacity of the tax agencies may be inadequate even if (1) and (2) are not problems. Such capacity can only be improved through training and technical assistance, which require long lead times. A clear road map is needed to guide actions in these areas over time and could be provided through surveillance. Where decisions under the control of the executive branch are the bottleneck, this can be taken up directly in program conditionality. When the constraint is the lack of legislation to empower tax agencies or implementation capacity that requires time to develop, surveillance should aim at evolving an agreed time frame for reform. This approach would allow conditionality in program situations to be more effectively focused on critical areas and would, therefore, be fully compatible with present streamlining initiatives.

Examples on the spending side

Programs often aim to contain the wage bill by capping public sector wages and/or reducing public employment by specified levels. However, progress in this area has been elusive. Short-term declines in real wages are usually followed by pressure for reversals, such as in the Romania and Ecuador programs. Although both the Tanzania and Costa Rica programs were able to achieve reductions in public sector employment, such progress is easily reversed after the program. Long-term civil service reform is therefore critical, but it is also problematic. Efforts to pass public employment legislation under a program have proven difficult. Attempts to pass legislation in the Costa Rica and Bulgaria programs were unsuccessful. Civil service reform initiatives require long preparation and consensus building. They should be encouraged in the context of longer-term programs such as EFFs and/or integrated, in close collaboration with the World Bank, into a longer-

[10]In considering these examples, we would like to reiterate that this report has focused on adjustment and policy reforms under specific IMF-supported programs and did not explore the links between past levels of TA and programs. In particular, FAD TA has been crucial for tracking HIPC spending and ROSC initiatives, areas not focused on in this evaluation.

[11]The problems encountered in improving the tax structure and strengthening tax administration over a long series of IMF-supported programs are discussed in more depth in a detailed case study of the Philippines prepared as part of the evaluation of prolonged use of IMF resources. See IEO (2002), Chapter 10, pp. 163–65.

Box 7.3. The Romania Program: Diminishing Returns to Raising Tax Rates

In order to meet quantitative targets in the Romania program, the statutory rate of social security contributions was increased from 35 percent in 1997 to 43 percent in 1998, reflecting both the deterioration in the finances of the public pension system and the establishment of the health social insurance fund. Budget revenue from social security contributions consequently surged from 7 percent of GDP in 1997 to 8.9 percent of GDP in 1998. However, the compliance rate was low—about 53 percent. Contribution arrears of large state-owned companies ballooned, with an increasing number of private companies following suit. As a result, total arrears to social security funds went up from 2.4 percent of GDP in 1997 to 3.4 percent of GDP in 1998.

Such circumstances do not warrant a further increase in contributions. It is thus rather surprising that the 1999 SBA relied on a hike of the statutory social contribution rate to the outstandingly high level of 60 percent. As a result, the compliance rate worsened in 1999 to about 44 percent, while arrears to social security funds increased to 3.8 percent of GDP. The private sector accounted for the bulk of the increase in contribution arrears, perhaps because state enterprises were closely monitored under the program.

Program projections implicitly incorporated a significant reduction in compliance rates. The revenue yield of social security contributions in 1999 was conservatively targeted at the same level as in 1998 (8.9 percent of GDP), including on account of the negative impact of the envisaged wage discipline upon the tax base. The actual yield was 10.7 percent of GDP. This revenue performance is partly explained by the fact that wage discipline was actually looser as compared with the targets of the program.

term framework of reforms specified under the broader road map discussed above.

Learning from the Past and the Role of Surveillance in Monitoring Progress

As argued in the previous chapter, programs often have too short a time frame to tackle major public finance reforms—particularly when programs are associated with crisis. Sustained reform, particularly in complex institutional areas calling for important political decisions, is better addressed in noncrisis years. The role of surveillance in setting a clear road map of structural reform and monitoring over time could greatly encourage this process.

In this chapter, we summarize our findings regarding learning from past experience and the role of surveillance in monitoring progress in structural reforms in the fiscal area and their links to programs. Specifically, we consider the extent to which (1) programs build on past reform efforts and try to learn from such efforts, (2) surveillance follows up and encourages reform, and (3) programs build on surveillance to address major public finance distortions (Appendix 8, Table A8.1 provides details). For each of these three areas, we explore a subset of questions as follows:

Learning from the past

- *To what extent do program documents analyze and evaluate past fiscal performance?*
- *To what extent do program documents specifically analyze and evaluate fiscal performance under the previous arrangement? Does self-standing surveillance (not associated with a program request or program review) tend to perform better in this area?*

Monitoring of fiscal reforms under surveillance

- *To what extent has surveillance flagged the need to accelerate fiscal reform in areas where implementation was lacking?*

Links between surveillance and programs

- *Were most major fiscal reform issues flagged during surveillance incorporated into the program?*
- *Were most problem areas taken up in programs identified by earlier surveillance?*

To address these questions, we reviewed surveillance activity over the three years prior to the program. This involved an analysis of 33 preprogram surveillance documents associated with the sample of 15 programs studied (Appendix 8, Table A8.2).

To obtain a quantitative estimate of the effectiveness of the IMF in each of these areas, the evaluation team's assessment in response to each question was classified into three categories: poor, mixed, and good performance. We then again constructed an index of performance by assigning a weight of 1, 0.5,

Box 7.4. Tax Reform in the 1996 Tanzania ESAF

The policy challenge

Tanzanian taxation prior to the 1996 program was characterized by far-reaching discretionary powers accorded the Minister of Finance, substantial statutory exemptions, including investment incentives, and widespread tax evasion. In 1994 discretionary tax exemptions amounted to the equivalent of over 20 percent of total recurrent revenue. The revenue losses and associated inequities were compounded by tax evasion.

The program, therefore, focused on improving tax administration through support and equipment to the newly created Tanzania Revenue Authority, curbing the discretionary powers of the Ministry of Finance in granting exemptions, diversifying the tax base, and establishing a tax appeals system. In parallel, energy sector fuel pricing and importation were liberalized.

Why revenue failed to increase

While the program's macroeconomic and trade reforms were relatively successful, progress on fiscal reforms was limited, with a serious gap between tax policy and implementation. The 1996 program aimed at revenue increases of 2 percent of GDP by 1999 while in practice total revenue fell by 2 percent of GDP. Some observers attribute the poor revenue performance to severe weather shocks (El Niño), the negative impact of the Asian crisis, and the contagion effects of the Great Lakes crisis. However, the economy grew at close to 4 percent a year during the program, higher than for many countries in the region. Thus shocks are not a sufficient explanation. More importantly:

- The program overestimated the speed at which institutional capacities could be strengthened, and VAT revenue projections were too optimistic. Tax evasion continued to be a serious problem. Adjusting tax legislation was important in modernizing tax administration, but much more attention should have been paid to capacity building. Lacking technical and managerial capacity, the Tanzania Revenue Authority was unable to implement the new policies expeditiously or to resist political pressure. Lack of technical competence and inadequate data on potential taxpayers led to poor tax assessments, inefficient coverage, and thus to revenue loss.
- While discretionary exemptions were largely eliminated at the government level, statutory exemptions for religious foundations, nongovernmental organizations, and other institutions remained substantial. In an earlier bid to attract investors, the government provided broad tax incentives to firms in mining and tourism. This minimized the revenue contribution of these growth sectors. Moreover, legal provisions for exemptions, most recently in the statutory provisions of the VAT, result in pressure to use them in ways not intended.

The policy sequencing also contributed to the revenue decline. In retrospect, tariffs were lowered too quickly before compensatory tax broadening measures, including strengthened administration, were in place. Increased corruption in the ports and customs administration were major contributors to the revenue decline. In the case of oil sector liberalization, the freeing of import licensing before setting up an industry regulator led to a situation of significant fuel smuggling. In addition, contrary to assumptions, lower tariff rates did not automatically increase tax compliance. Also, the tax base was eroded due to the failure of several inefficient industrial enterprises.

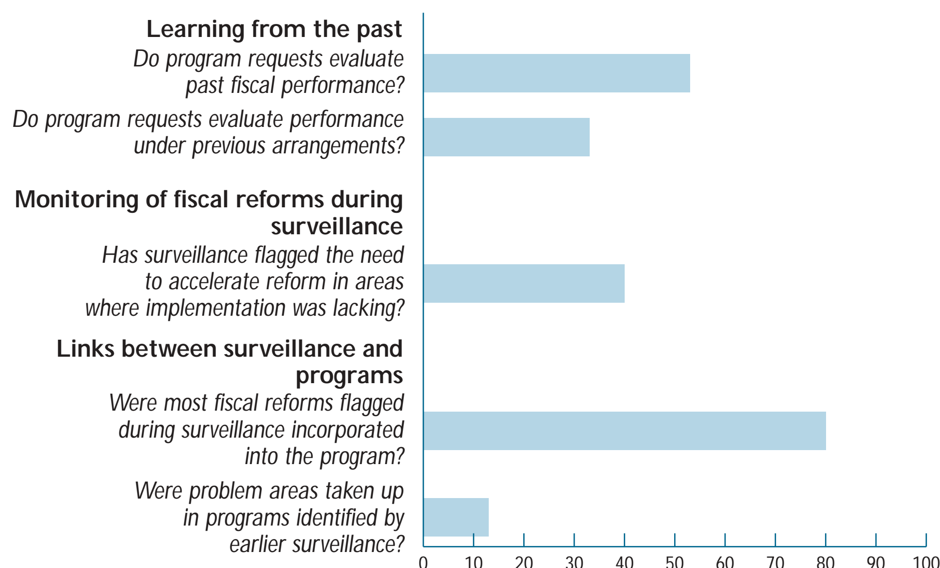
and 0, respectively, for "good," "mixed," and "poor" performance.

Figure 7.3 summarizes the resulting average performance in each of the five areas. "Learning from the past" appears as an area of generally poor results. Program requests are only partly successful in evaluating past fiscal performance—with an index of success of about 50 percent. The results are worse (35 percent success) when documents are judged on a more pointed question: how well they analyze performance and policy failures under the previous arrangement. Overall, programs tend to focus on performance during the previous year and rather independently of previous arrangements. Few efforts are made to analyze the factors behind past policy failures.

Efforts during surveillance to *flag the need to accelerate reforms* are also limited, with an index of success of about 40 percent.

Finally, Figure 7.3 shows a sharp asymmetric link between the issues identified under surveillance and those taken up by the subsequent IMF-supported program. Problem areas flagged under previous surveillance are typically incorporated fairly well in programs and this is the area of best performance (80 percent). On the other hand, programs include many reform areas that were not flagged early on by surveillance. In fact, this is by far the worst area of performance (performance is good in only 15 percent of cases). Although unexpected developments and shocks may call for programs to include fiscal reforms not previously flagged in surveillance, we would expect this to be the exception rather than the rule. Moreover, introducing issues which have not previously been flagged as a concern may reduce country ownership and suggest that the IMF is using its leverage to push for reforms that are not essential (since they were not previously flagged).

Figure 7.3. Index of Performance: Learning, Follow-Up, and Links Between Programs and Surveillance
(In percent)



Source: IEO staff calculations based on program documents.

Inevitably, the average value of the index masks significant variation. To explore this variability, Figure 7.4 shows the distribution of cases behind the averages. We also give examples of specific cases to provide a better sense of such variability and identify best practice.

Learning from the past

Program requests show a satisfactory analysis of past fiscal performance in only 40 percent of cases and in one-third of cases that analysis is poor. These results deteriorate when programs are judged by how well they examine performance under the last arrangement; only one-quarter perform well and almost two-thirds perform poorly.

There are, however, good examples, where program requests perform well on both counts, such as the Algeria, Philippines, and Senegal programs (see Box 7.5).

An attempt was made to assess whether more recent program-request documents make stronger efforts to evaluate past fiscal performance than was done in the 15 programs originally studied. Eight of the 15 countries had more recent programs, namely, Algeria, Bulgaria, Jordan, Pakistan, Peru, Romania, Tanzania, and Uruguay.[12] These 8 programs were then examined under the same criteria, namely the extent to which they evaluated and analyzed past fiscal performance. The results show no significant improvement with respect to the earlier results—only 4 of the 8 cases (the Algeria, Tanzania, Jordan, and Romania program-request documents) were judged as successful in this area.

In order to assess whether performance is better for self-standing surveillance reports—which potentially have the opportunity to analyze progress and learn lessons without program distractions and operational pressures—we looked specifically at six free-standing Article IV documents that preceded the program being studied.[13] Although the sample is small, the results are revealing. Only the Romania 1998 Article IV conducted an in-depth examination of the main fiscal issues of the prior 18 months and the reasons why the previous arrangement went off track.

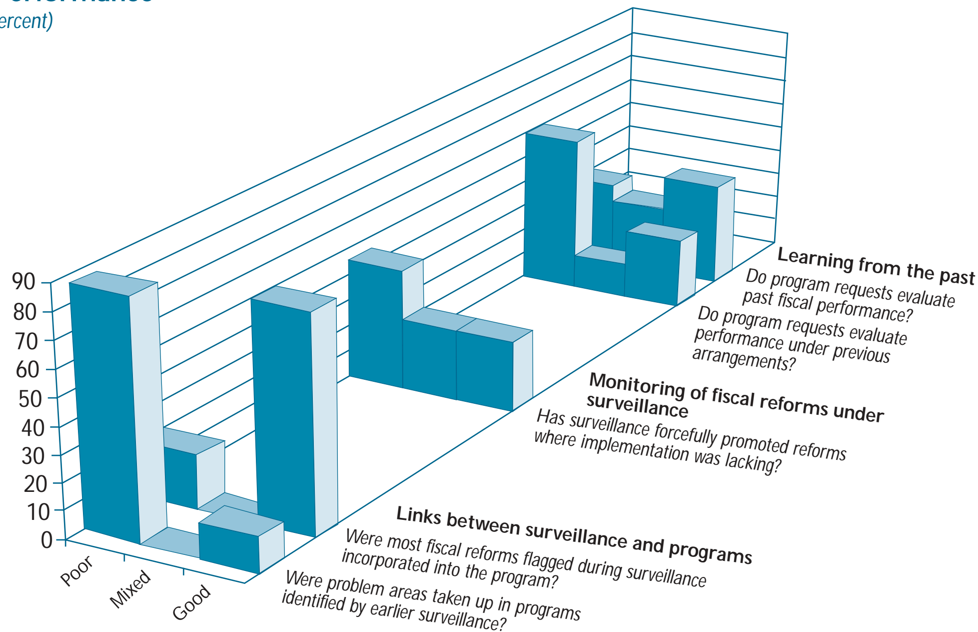
Monitoring of reforms under surveillance

In only one-quarter of cases was surveillance forceful in flagging the need for reform where im-

[12]The specific program requests examined are Algeria EFF 1995, Bulgaria SBA 2002, Jordan SBA 2002, Pakistan PRGF 2001, Peru SBA 2002, Romania SBA 2001, Tanzania PRGF 2000, and Uruguay SBA 2002.

[13]These six Article IV documents are a subset of the 11 self-standing Article IV documents included in the sample. This subset of Article IV includes Algeria 1992, Costa Rica 1994, Ecuador 1997, Romania 1998, Tanzania 1995, and Venezuela 1993.

Figure 7.4. Distribution of Cases According to Performance
(In percent)



Source: IEO staff calculations based on program documents.

plementation was lacking. In 40 percent of cases this effort was weak.

Ukraine is one of the better cases. Both Article IV consultations (for 1995 and 1997) thoroughly identified and analyzed reasons for failure in past reform implementation, as well as remaining implementation risks. The 1995 report identifies and discusses four areas most affected by slippages in implementation of the 1994/95 stabilization program (SM/95/320). Separate sections discuss the problem of external arrears and the social safety net. Implementation issues are explicitly analyzed and specific measures recommended. The 1997 consultation identifies the main risks to the program and singles out risks to fiscal policy and the budget.

The 1995 Egypt Article IV report was candid in focusing on areas of disagreement between the staff and the authorities on such issues as wage bill reduction, public sector employment cuts, civil service reform, privatization, and social safety net issues where the staff pressed for improved targeting of social transfers and less reliance on generalized subsidies.

The surveillance process in Bulgaria is a good example of improvement over time. The 1995 Article IV is weak on recommendations. Vague statements were made such as: "the staff underscored the importance of slowing wage increases as much as was feasible" [but with no target] or "the staff recommended that the authorities focus on expenditure rationalization rather than spending cuts" [with no specifics]. In contrast, the 1997 Bulgaria Article IV is clear on recommendations and evaluation of progress (or lack of it), with structural reforms. The report contains a template with the status of conditions for completion of the first review under the SBA. It discusses special efforts to accelerate banking system reform, and follow-up on the efforts of the new government to implement legal wage limits. Nevertheless, the report has very little on expenditure control, with the only reference being: "the authorities are also taking steps to improve expenditure control by including in the budget provisions to limit commitments of spending agencies to 90 percent of allocated funds."

On the other hand, preprogram surveillance in Peru was not forceful in encouraging reforms to reduce tax evasion, contain the growth of the public sector wage bill, and increase social spending. Another example is the 1997 Article IV for Ecuador—the only surveillance exercise in the 1996–99 period. The report failed to highlight the dramatic deterioration of the banking sector and the need for urgent actions particularly in the supervision area. Instead the recommendations are buried in the middle of the report rather than being flagged up front in the summary. Little attention was given to documenting and addressing the massive tax evasion taking place—

Box 7.5. Good Examples of Learning from the Past

The Algeria report has three well-focused chapters that evaluate past economic performance. This includes reform implementation during 1989–91 (including the past two IMF arrangements), and developments from 1992 to 1994. The report also discusses major fiscal distortions not raised before. It proposes new policy recommendations to encourage reform (on government investment spending, wage policy, and various revenue measures). For example, the report proposes that government investment should be limited to priority projects and proposes transferring investment financing responsibility from the Treasury to the enterprises and banking system.

The Philippines program-request document, which was prepared jointly with an Article IV surveillance report, thoroughly discusses performance under the previous EFF. The assessment includes main goals, achievements, and policy failures for the overall program and fiscal policy. Box 2 on "The Extended Arrangement in Retrospect" provides a brief and clear summary of the main areas of progress as well as lack of progress in fiscal policy and reforms. Box 4 on the "Comprehensive Tax Reform Package" discusses the main elements of the reform and the main implementation issues.

The Senegal report does a good job overall, although there is scope for more specificity and a more analytical look at the past. The executive summary and the chapter "Performance Under the Previous ESAF-Supported Program and Recent Developments" comprehensively evaluates fiscal performance under the previous arrangement. The document includes a summary of selected policy performance indicators with the main achievements in fiscal policy during the three years of the previous arrangement. There is a thorough discussion of why some reform implementation was behind schedule (energy and privatization). The main fiscal achievements under the previous arrangement are clearly addressed. However, fiscal targets and objectives for the previous ESAF are not made explicit and are discussed only for the previous year.

most of the references being focused on the need to increase VAT rates.

Links between surveillance and IMF-supported programs

Programs successfully include issues identified during surveillance. On the other hand, surveillance fails to identify many of the reforms that subsequent programs found necessary to incorporate.

Programs incorporate the main issues flagged during surveillance in about 80 percent of cases. For example, the Tanzania program reiterates many of the issues raised during surveillance, notably the need to strengthen the implementation capacity of public sector institutions. During surveillance, inadequate institutional capacities were held responsible for poor tax administration, tax evasion, and inability to formulate and implement policies. Similarly, the Uruguay program request explicitly targets areas flagged during surveillance, including the need to continue improving tax administration and tax compliance. It also clearly and forcefully recalls measures raised in earlier surveillance reports to strengthen public sector banks, restrain wages, and complement social security system reform.

Preprogram surveillance fails to identify problem areas dealt with by programs in almost 90 percent of cases. Only in the Pakistan and Philippines programs were almost all issues under the program previously identified in surveillance.

In the case of Bulgaria, preprogram surveillance failed to flag important measures to enhance fiscal transparency (such as explicitly incorporating into the budget quasi-fiscal costs of restructuring and liquidating state-owned enterprises and any support provided to them), and the consolidation of the largest extrabudgetary funds into the budget. Surveillance also did not flag tax administration measures to enhance tax collection (such as the development of a tax collection strategy, including enforcement, audit, and fraud investigation), measures to improve expenditure controls, and, finally, measures to rationalize and increase the cost-effectiveness in the supply of public health services, all of which were included in the subsequent program.

Surveillance prior to the Peru program did not address the need to rationalize employment in the public sector and reform the public sector wage structure.

Conclusions

Focusing on the unfinished reform agenda and reducing vulnerabilities to future crises will require strong follow-up during surveillance as well as continuity in successive programs. At present, surveillance

does not forcefully flag policy inaction—many times it is insufficiently candid in language. Many program-request documents are insufficiently linked to past outcomes and past reform attempts. Although based on a very small sample, self-standing surveillance does not seem to yield better results. This is a missed opportunity because we would expect that surveillance not associated with a program request or review would have a genuine opportunity to take a more strategic perspective on both assessing whether fiscal reforms over time add cumulatively to better fiscal systems, and spelling out clearly what the remaining fiscal agenda for the future should be.

Surveillance should play a much more forceful role in providing a medium-term road map of structural reforms to be followed up over time, with or without programs. Progress and reasons for inaction should be reported candidly. That road map would then provide guidance for the specific reform priorities to be taken up in successive programs—this being particularly important in repeat users of IMF resources. Such an analysis would also provide the broader strategic overview of fiscal reform priorities as well as the success or failures of past efforts that would help inform choices about the priorities for reform implementation in any future programs.