



Working Paper

47

WHAT CAN WE KNOW ABOUT THE EFFECTS OF IMF PROGRAMMES?

**Tony Killick, Moazzam Malik
and Marcus Manuel**

**Results of ODI research presented in preliminary form
for discussion and critical comment**

ODI Working Papers available at September 1990

- 24: Industrialisation in Sub-Saharan Africa: Country case study: Cameroon**
Igor Karmiloff, 1988, £3.00, ISBN 0 85003 112 5
- 25: Industrialisation in Sub-Saharan Africa: Country case study: Zimbabwe**
Roger Riddell, 1988, £3.00, ISBN 0 85003 113 3
- 26: Industrialisation in Sub-Saharan Africa: Country case study: Zambia**
Igor Karmiloff, 1988, £3.00, ISBN 0 85003 114 1
- 27: European Community Trade Barriers to Tropical Agricultural Products**
Michael Davenport, 1988, £4.00, ISBN 0 85003 117 6
- 28: Trade and Financing Strategies for the New NICs: the Peru Case Study**
Jurgen Schuldt L., 1988, £3.00, ISBN 0 85003 118 4
- 29: The Control of Money Supply in Developing Countries: China, 1949-1988**
Anita Santorum, 1989, £3.00, ISBN 0 85003 122 2
- 30: Monetary Policy Effectiveness in Côte d'Ivoire**
Christopher E Lane, 1990, £3.00, ISBN 0 85003 125 7
- 31: Economic Development and the Adaptive Economy**
Tony Killick, 1990, £3.50, ISBN 0 85003 126 5
- 32: Principles of policy for the Adaptive Economy**
Tony Killick, 1990, £3.50, ISBN 0 85003 127 3
- 33: Exchange Rates and Structural Adjustment**
Tony Killick, 1990, £3.50, ISBN 0 85003 128 1
- 34: Markets and Governments in Agricultural and Industrial Adjustment**
Tony Killick, 1990, £3.50, ISBN 0 85003 129 X
- 35: Financial Sector Policies in the Adaptive Economy**
Tony Killick, 1990, £3.50, ISBN 0 85003 131 1
- 37: Judging Success: Evaluating NGO Income-Generating Projects**
Roger Riddell, 1990, £3.50, ISBN 0 85003 133 8
- 38: ACP Export Diversification: Non-Traditional Exports from Zimbabwe**
Roger Riddell, 1990, £3.50, ISBN 0 85003 134 6
- 39: Monetary Policy in Kenya, 1967-88**
Tony Killick and F M Mwega, 1990, £3.50, ISBN 0 85003 135 4
- 40: ACP Export Diversification: Jamaica, Kenya and Ethiopia**
Christopher Stevens, 1990, £3.50, ISBN 0 85003 136 2
- 41: ACP Export Diversification: The Case of Mauritius**
Mathew McQueen, 1990, £3.50, ISBN 0 85003 137 0
- 42: An Econometric Study of Selected Monetary Policy Issues in Kenya**
F. M. Mwega, 1990, £3.50, ISBN 0 85003 142 7
- 43: Differences in Econometric Performance between Franc Zone and other sub-Saharan African Countries**
Christopher E. Lane and Sheila Page, 1991, £3.50, ISBN 085003 148 6
- 44: Monetary Policy Effectiveness in Indonesia, 1974-1990**
Christopher E. Lane, David C. Cole and Betty F. Slade, 1991, £3.50, ISBN 0 85003 150 8
- 45: Monetary Control in Ghana: 1957-1988**
Nii Kwaku Sowa, 1991, £3.50, ISBN 0 85003 152 4
- 46: The IMF in the 1990s: Forward to the Past or Back to the Future?**
Graham Bird, 1991, £3.50, ISBN 0 85003 153 2
- 47: What Can We Know About the Effects of IMF Programmes?**
Tony Killick, Moazzam Malik and Marcus Manuel, 1991, £3.50, ISBN 0 85003 154 0
- 48: Country Experiences with IMF Programmes in the 1980s**
Tony Killick with Moazzam Malik, 1991, £3.50, ISBN 0 85003 155 9

00002827

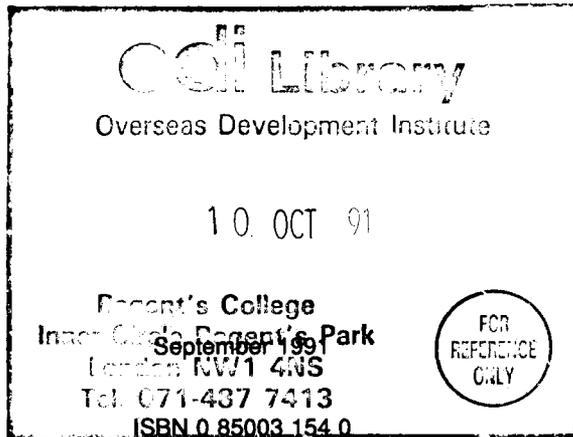


Overseas Development Institute

WORKING PAPER 47

**WHAT CAN WE KNOW ABOUT THE EFFECTS
OF IMF PROGRAMMES?**

Tony Killick, Moazzam Malik and Marcus Manuel



OVERSEAS DEVELOPMENT INSTITUTE
Regent's College
Inner Circle, Regent's Park
LONDON NW1 4NS

Preface and Acknowledgements

ODI Working Papers present in preliminary form work resulting from research under the auspices of the Institute. Views expressed are those of the authors and do not necessarily reflect the views of ODI. Comments are welcomed, and should be addressed directly to the authors.

This paper is one of a series of drafts for a study currently under preparation at ODI by Graham Bird and Tony Killick with the provisional title of *The IMF and Developing Countries: Its Role in the 1990s*. The completed report will review developments in the 1980s; examine the Fund as a source of finance and issues in its lending policies; review the theory and practice of IMF policy conditionality and of heterodox alternatives to it; and explore the future role of the Fund. The following titles are published contemporaneously with this paper and others will follow later:

- 46 The IMF in the 1990s: Forward to the Past or Back to the Future?
Graham Bird
- 48 Country Experiences with IMF Programmes in the 1980s
Tony Killick with Moazzam Malik

Tony Killick is Senior Research Fellow at the Overseas Development Institute. At the time of preparing this paper Moazzam Malik and Marcus Manuel were Research Assistants at ODI. The project under which this Working Paper has been prepared is funded by the Overseas Development Administration, whose support is gratefully acknowledged. Neither they nor our respective employers necessarily agree with the contents of this Working Paper, which is our responsibility alone.

CONTENTS

| | <u>Page</u> |
|--------------------------------------------------|-------------|
| INTRODUCTION | 1 |
| I. THE PRESENT STATE OF KNOWLEDGE | 2 |
| I. 1 Methodological pitfalls | 2 |
| I. 2 A menu-of-information approach | 5 |
| I. 3 The menu of information | 8 |
| I. 4 Sources of difficulty | 16 |
| II. A STUDY OF PROGRAMME NON-COMPLETIONS | 20 |
| II. 1 Explanation and justification | 20 |
| II. 2 The results | 22 |
| III. A BEFORE-AFTER ANALYSIS OF THE 1980s | 28 |
| III.1 Nature of the study | 28 |
| III.2 The results | 28 |
| METHODOLOGICAL APPENDIX | 36 |
| REFERENCES | 39 |

INTRODUCTION ¹

Controversy about the appropriateness and effects of the stabilisation programmes associated with the IMF has a venerable tradition in international economics and has generated a substantial literature. Both the Fund and its critics like to believe that its programmes² have powerful effects, although they disagree over whether these are favourable or malign. Others have suggested that the controversy is a case of much ado about nothing: that the effects are relatively slight in either direction.

This uncertainty is unsatisfactory. At any one time a large number of developing countries have IMF programmes in place, in the negotiation of which large amounts of scarce skills and information have been deployed by both parties. The macroeconomic policies of borrowing countries are strongly influenced, not to say dominated, by the terms of the programme. Virtually all the Fund's financial resources and a large proportion of its staff are devoted to these programmes. Debt reschedulings are usually conditional on the debtor government signing a programme agreement with the Fund. A good many credit and investment decisions, by both public and private agencies, are similarly influenced by whether or not a Fund programme is in place. It would be absurd to place so much weight on the existence of an IMF programme if, in fact, they had little effect.

At the same time, as we will see, there are major difficulties in the way of forming a definitive judgement on programme effects. This Working Paper aims to make a contribution to the debate in two ways. Part I surveys the present state of the literature which applies quantitative tests to this issue, arguing that we can deduce from it more about programme effects than is sometimes suggested. Parts II and III present the results of new research which throws additional light on programme impact.

This Working Paper should be read in conjunction with a companion piece³ which presents the results of a survey of case studies, intended to take advantage of the greater depth permitted by the case study approach while overcoming the disadvantage that one cannot derive general conclusions from individual studies. That Working Paper will also attempt to draw together the content of both in order to offer a summing up on the present state of knowledge on the effects of IMF programmes in developing countries.

¹ The authors are grateful for helpful comments by Margaret de Vries, Gerry Helleiner, Mohsin Khan and Jacques Polak, although none of them is to be held responsible for remaining shortcomings.

² Although it is useful shorthand to refer to 'IMF programmes', the Fund always refers to 'Fund-supported' programmes in order to convey that the programmes are the property of the governments which sign them. The extent to which this is truly the case is, of course, one of the points of controversy in the literature. Our shorthand use of 'IMF programmes' is not intended to prejudice this issue!

³ Tony Killick with Moazzam Malik, 'Developing Country Experiences with IMF Programmes in the 1980s', *ODI Working Paper No. 48*, September 1991.

I. THE PRESENT STATE OF KNOWLEDGE

I.1 Methodological pitfalls

There are formidable obstacles in the way of arriving at firm evidence on the effects of IMF programmes. Thus:

- (a) A large number of influences bear upon the performance of an economy and it is difficult in practice to disentangle the effects of these other factors from those of an IMF programme.
- (b) It is difficult to distinguish between the effects of the policies in Fund programmes and the effects of the Fund (and other) finance which accompanies the programmes. This may be a particularly large difficulty with credits linked to World Bank structural adjustment lending, *e.g.* programmes under the Enlarged Structural Adjustment Facility (ESAF), because large volumes of finance are sometimes associated with such policy packages. In such cases evaluating the effects of the programme of either one institution is rather artificial and it may be preferable to study the joint effects of them both. However, this has not been attempted in the literature. Similar problems arise where Fund credits are linked to major debt relief operations, for debt relief is equivalent to additional finance.
- (c) There are difficulties in selecting adequate performance indicators in statistically-based cross-section analyses. This is perhaps most acute in the case of the balance of payments (BoP), the improvement of which is the chief objective of IMF programmes. The overall balance (*i.e.* the balance on monetary account with opposite sign) is the most commonly used indicator but this is vulnerable to the difficulties outlined in (a) above and to the influence of capital inflows and debt relief triggered by agreement with the Fund. The current account balance is another popular indicator but the difficulty with this is that governments facing a foreign exchange crisis with exhausted reserves have to find ways of limiting the current account deficit to whatever money is expected to be available to finance it. They often do so by severe import cuts - which are liable to have adverse effects on economic performance, including ability to export⁴, so that a reduction in the current account deficit does not necessarily tell us anything about the underlying strength of the BoP situation. Use of the overall balance is also affected by this limitation. The essential difficulty is that progress on the BoP cannot adequately be monitored by reference to one or two residual balances. To a lesser extent, similar problems arise with other performance indicators commonly used.

⁴ Khan and Knight, 1988, tested for the influence of import availability on the export performance of 34 developing countries and found a large and highly significant positive correlation between them.

- (d) There are problems also with the period of analysis, with results often sensitive to the choice made. One issue here is whether the impact of Fund programmes should be assessed only for the period of the programmes or should extend beyond them. The time-lags between changes in policy variables such as the exchange rate or domestic credit and BoP results suggest evaluation over a more-or-less extended period but the longer this is extended the louder becomes the extraneous 'noise' in the tests. The period of analysis also creates difficulties in comparisons of economic performance in countries with and without IMF programmes: given that programmes are agreed at different times for varying periods, it is not clear over precisely what period the performance of the non-programme countries should be measured.⁵
- (e) The rigour of policy conditionality varies according to programme type. Only rather relaxed conditionality is attached to first-tranche stand-by credits, as compared with upper-tranche credits. ESAF conditionality is said to be the most demanding of all.⁶ Even among programmes of the same type there are considerable variations, depending on the initial situation, the government's past record in dealings with the Fund, and the exercise of influence on the Fund management by interested major-shareholder governments. Presumably results should differ in extent and speed according to the severity of conditionality and the level of supporting finance but multi-country tests generally treat all programmes as equal.
- (f) Finally, some programmes are more fully implemented than others and effectiveness tests should ideally adjust for degree of implementation. Indeed, as we will see in Part II, a large proportion of all programmes break down during their intended lifetime.

Above all, however, the literature on IMF programme effects is dominated by the problem of the counterfactual, taking as the central question, do Fund programmes result in a better situation than would have obtained in the absence the programmes? Khan, whose 1990 article provides an up-to-date survey of the literature and a substantial contribution to methodology and evidence, puts the point eloquently:

The counterfactual is perhaps the most appealing yardstick against which to assess program performance and the standard most widely employed in economics to measure the impact of government policy interventions. What would have happened in the absence of a Fund-supported program is by no means the only standard against which to judge the outcome of programs, but in many cases it is the most appropriate one. However, the counterfactual cannot, by definition, be observed and must be estimated or

⁵ For an example of this problem, and of the sensitivity of results to the period chosen, see Gylfason, 1987.

⁶ Goreux, 1989, p.150.

approximated. The various approaches used in evaluation studies should thus be judged in terms of how good they are in providing estimates of the counterfactual. [Our emphasis].

There are perhaps three, but linked, principal reasons why it is important to use the counterfactual yardstick when enquiring into the effects of Fund programmes. The first is the importance of disentangling the (generally adverse) effects of the situation of the economy - commonly a crisis situation - immediately prior to adoption of a programme from the effects of the programme itself. In the past some critics of the Fund have failed to attempt this and have illegitimately imputed to the Fund adverse developments in the economy which might more appropriately be attributed to the initial situation. If the economy starts with a foreign exchange crisis, rapid inflation, major supply-side bottlenecks and large-scale excess demand it is scarcely surprising if in the following two or three years there is little growth, consumption standards fall and investment declines.

This tendency to attribute to programmes an economic deterioration which is, in fact, pre-determined by the initial situation is compounded by the fact that by far the easiest, and most commonly employed, type of cross-country test to apply in empirical investigations of programme effects is the comparison of economic variables before and after programme introduction - the so-called before-after test.

A second reason why it would be desirable to apply a counterfactual test is that programmes are commonly knocked off course by shocks which are beyond the control of governments and the Fund. Abnormal weather, organised violence, changing world capital market conditions are examples but unexpected changes in borrowing countries' terms of trade are probably the most common and important. Given the turbulence in world economic conditions in the 1970s and 1980s, it was inevitable that a substantial proportion of Fund programmes would 'fail' because of such factors.⁷ However, the counterfactual, by definition, would be no less affected by these exogenous shocks so that comparisons between this and programme results would eliminate the bias.

This leads to a third reason why the counterfactual problem is regarded as so central in evaluating Fund programmes: that this literature has a strongly normative content, for it is inevitably read as evidence on the effectiveness of the IMF as an institution. This is unfortunate. Since both exogenous shocks and non-implementation of programme provisions by borrowing governments are reasons for programme ineffectiveness which cannot be laid at the door of the Fund, it would be desirable to separate the 'positive' assessment of programme effects from normative evaluations of the success of the IMF - and information about the counterfactual would considerably facilitate this separation.

⁷ For a discussion of this and other reasons for programme 'failure' in the 1970s see Killick *et al.*, 1984A, pp.250-64.

I.2 A menu-of-information approach

How should we respond to these difficulties, particularly to the unknowability of the counterfactual? Although we have shown that the problem of the counterfactual must be taken seriously, we suggest now that it has come to exert an undue dominance in this literature. In particular, it seems wrong to use their adequacy in proxying the counterfactual as a decisive test of alternative techniques for testing programme effects for, as Khan noted, other criteria are relevant too.

Take the before-after tests, for example. They do not cope well with the counterfactual problem - but they do still provide useful 'positive' information. It is perfectly sensible to ask, did this programme improve on the initial situation? It was almost certainly in order to do so that the government negotiated a Fund credit! The Fund's articles state that its credits are intended to give governments an opportunity to 'correct maladjustments in their balance of payments', and this implies that programmes are intended to improve on the initial BoP situation. To the extent that they do not, we then have to search for the reasons for the shortfalls. These may or may not lie within the responsibility of the IMF or the borrowing government but it is certainly useful information to know whether the situation was improved.

More broadly, the position taken here is that the various tests of programme effects each provide useful, although limited, information and we should therefore use them all in order to build up a menu of 'positive' information. We must, however, be careful about using these results, and mindful of the pitfalls, for any normative assessment of Fund effectiveness. In this spirit, what follows is a brief resumé of the positive information that the various available tests can yield, and of their chief weaknesses, summarised in Table 1. Using this approach we can build up an inventory of what we can (and cannot) know about programme effects, as a preliminary to a summary of what the available evidence tells us on this subject.⁸

□ *Before-after tests.*⁹

Discussed above. These have the advantage of being relatively easy to conduct. They yield information on whether programmes were associated with an improvement on initial situation. Because they are relatively easy, they can more readily be used also to give information on whether programme effects are sustained, and whether results are affected by country or programme type. Results can be misleading, however, because of exogenous shocks and are easily mis-used in normative discussions.

⁸ For a more extended discussion of the methodological issues, although excessively dominated by concerns about the counterfactual, see especially Goldstein and Montiel, 1986. Also Khan, 1990.

⁹ A substantial number of such researchers have used this method. Among the principal reports of research findings are: Reichmann and Stullson, 1978; Connors, 1979; Killick *et al.*, 1984A; Goldstein and Montiel, 1986; and Khan, 1990.

□ *Target-actual tests:*¹⁰

These compare actual outcomes with the targets written into Fund programmes. They therefore give us an indication of the extent to which programmes achieve intended results. However, such quantified targets are no better than the models which generate them in simulations, so that discrepancies between targets and actualities may be a result of poorly-performing models rather than intrinsically unsatisfactory economic results. Like the before-after approach, this does not provide a proxy for the counterfactual and yields no information on the costs of achieving results. It might give unduly negative results because of unrealism in programme targets - but that the targets may be unrealistic is itself useful information. Its main operational snag is that it requires information on programme targets which is not publicly available. For multi-country studies its use is confined to IMF staff or researchers to whom the Fund is willing to release the information.¹¹

□ *With-without tests:*¹²

These compare a sample of programme countries with a control group of other countries which did not adopt Fund programmes. In this way it seeks to tackle the counterfactual question, making the experiences of the control group a proxy for what would have happened in the programme countries if they had not gone to the Fund. The great difficulty with this approach is to select a truly comparable control group. Although later attempts (*e.g.* Goldstein and Montiel, 1986; Gylfason, 1987) make major efforts to ensure this, it is intrinsically difficult because the decision by control countries with equally severe BoP problems not to adopt a Fund programme can be expected to be of wide significance, vitiating the comparability of the two samples.¹³

¹⁰ See Reichmann, 1978; and Zulu and Nsouli, 1985. The results of some unpublished IMF tests are reported in Killick *et al.*, 1984A, chapter 7. See also Edwards, 1989. Heller *et al.*, 1988, also contains target-actual information.

¹¹ The only such example of which we are aware is provided by Edwards, 1989, who obtained information on programmes in 1983-85 - see his Table 5.

¹² For examples see Donovan, 1982; Loxley, 1984; Goldstein and Montiel, 1986; and Gylfason, 1987.

¹³ Thus, Gylfason, 1987, went to considerable lengths to ensure that his programme and non-programme countries had BoP difficulties of equivalent severity. However, it is likely that some of his control group did not go to the IMF on ideological grounds, or due to incompetence, or because they gave low priority to macroeconomic management. This group could thus be expected to do poorly in macroeconomic terms, so that their experiences were not a good proxy for what would have happened in the programme countries in the absence of the Fund. In general, the contrasts would impute unduly positive results to the programmes. Examination of the composition of Gylfason's control group reveals

□ *Comparison-of-simulations:*¹⁴

These are based on policy models of the behaviour of a single economy, or group of economies. These can thus be used to predict outcomes when IMF-type policies are introduced and to compare with the counterfactual situation. Depending on model specification, they can also adjust for the effects of exogenous variables. Their usefulness depends on the appropriateness of the model used, however, and their danger is that the results reflect model specification rather than reality. Different models could produce substantially varying results. Also, this technique does not yield direct information on programme effects, *per se*, and it assumes that model parameter values would be unchanged by the policies adopted.

□ *Generalised evaluation:*¹⁵

This proceeds by postulating a model in which various policy and other explanatory variables are incorporated in regression equations which take the performance of the BoP, inflation and economic growth as dependent variables and in which a dummy variable is included to catch the influence of any Fund programme. This approach thus offers a direct test of programme impact, although its inclusion simply as a dummy means that it cannot provide refined information about programme effects. Here again, the method is as good as, and no better than, the model employed.

□ *Extent of programme completion:*

This provides information on the extent of implementation of programme provisions because governments are refused access to credit tranches when they fail to comply with the 'performance criteria' written into the programmes. As such, it throws light on an area which most of the other tests neglect: the reasons for programme shortfalls. Once the basic information has been collected, it can readily be classified so as to explore results for different programmes and country types. Part II below sets out the case for the usefulness of this approach and reports the results of such a test - the only systematic example of which we are aware.

this to be very heterogenous, and large dispersions around the mean values of the variables used by him.

¹⁴ I have borrowed this expression from Khan, 1990. There have been various uses of this method in single-country case studies. The chief examples of use in cross-section analyses are Khan and Knight, 1981 and 1985.

¹⁵ Again, I have borrowed from Khan, 1990, which is also the chief example of the use of this technique. See also Goldstein and Montiel, 1986.

□ *Country or programme case studies:*

In principle, the case study approach can avoid the various disadvantages adduced above and can go much further in explaining the reasons behind the results obtained. Its abiding disadvantage is that it does not yield results that can be generalised across programmes or countries. Indeed, it can be rather destructive of the desire to arrive at general assessments by emphasising the uniqueness of each case. A partial solution can be provided by undertaking a number of co-ordinated case studies but this is an approach which is particularly demanding of resources.¹⁶ An alternative attempt to overcome these limitations is made in Working Paper 48, which is a comparative study of the results of a large number of already-published individual case studies, although it has to cope with the difficulty that these were not devised within a common analytical framework.

An attempt is made in Table 1 to summarise the types of information that the various methods of evaluation are able to yield, and their chief weaknesses. While some are more informative than others, each method has its uses and by combining them we ought to be able to build up a substantial menu of 'positive' information about programme effects, which is the next task. Before moving to this, however, we should note the inability of most multi-country studies to go far in explaining the results obtained, as indicated in the lower part of Table 1 (items 8-11). If we are concerned to understand why programmes are sometimes unable to achieve their objectives, the table suggests four lines of explanation: the influence of the initial situation and of exogenous shocks; incomplete programme execution; inadequate programme impact on the intended instrument variables; and inadequate financial support for the programmes. To these can be added a fifth: possible defects in programme design. While certain of the multi-country approaches can attempt to screen out the effects of the initial situation and some exogenous shocks, they throw little light on the other factors. This is an area in which the case study approach scores especially heavily.

I.3 The menu of information

We are now ready to return to our original questions, what can we know? Given the varying uses of the tests just described, what do the results of the published studies cited in the footnotes tell us about the effects of Fund programmes? In approaching this question we will concentrate mainly on the statistically significant results of the various studies, but not to the exclusion of other suggestive results which may fail the 95% minimum significance level or for which statistical significance was not calculated. No

¹⁶ This technique was used in an earlier ODI study, with the case studies reported in Killick *et al.*, 1984B, whose results are utilised in the overview report, Killick *et al.*, 1984A. The WIDER country studies whose results are summarised in Taylor, 1988, also use this method, although this project does not attempt to measure the effects of IMF programmes *per se*. See also Heller *et al.*, 1988.

Table 1:

Uses and limitations of alternative tests of IMF programmes

| | <i>Before - after</i> | <i>Target - actual</i> | <i>With - without</i> | <i>Comparison of simulations</i> | <i>Generalised evaluation</i> | <i>Completed - uncompleted</i> | <i>Case- study</i> |
|------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|---------------------------|--------------------------------------|-----------------------------------|------------------------------------|------------------------|
| <u>Programme results</u> | | | | | | | |
| 1. Does the programme improve on the initial situation? | Y | Y | Y | | Y | | Y |
| 2. Do programme countries do better than non-programme? | | | Y | Y | Y | | |
| 3. Do programmes improve upon likely alternative outcomes? | W | W | D | Y | Y | W | Y |
| 4. Are programme effects sustained? | Y | | Y | | Y | Y | Y |
| 5. At what costs are their results secured? | | W | | | Y | | Y |
| 6. Can the results be generalised? | | | D | D | Y | | W |
| 7. Do results differ for: | | | | | | | |
| (a) country types? | Y | | Y | | | Y | |
| (b) programme types? | Y | | Y | | | Y | |
| <u>Result determinants</u> | | | | | | | |
| 8. Influence of exogenous factors | W | | | Y | Y | | Y |
| 9. Programme implementation | | | | | W | Y | Y |
| 10. Impact on instrument variables | | Y | | | | | Y |
| 11. Impact on financial flows | | | | | | | Y |
| Key: | Y = The test provides useful information bearing on the issue in the left-hand column; W = The test is particularly weak in this area; D = Debatable. There is disagreement in the literature on whether this test can provide useful information of the type asked for. | | | | | | |

attempt will be made to summarise systematically the results of the various studies.¹⁷ Our purpose is rather to build up a description of what information can reasonably be inferred about the effects of Fund programmes. Since we wish this to be relevant to the contemporary situation, special weight will be attached to tests which include data for the 1980s. The discussion will follow the sequence set out in Table 1. However, we will not follow the customary procedure of treating GDP growth as a target variable, bringing it in rather in the discussion of the costs of adjustment.¹⁸ We take the BoP indicators as the chief target variables, with reductions in the inflation rate as a second-order goal.

Our first question, then, is whether Fund programmes result in an improvement on the initial situation. The results of before-after tests are the most pertinent here. Generally speaking, the results of these are not encouraging. None of them shows a significant improvement in the BoP current account. The results are more mixed on the overall balance but of the general evaluations, only the most recent, utilising data for 1973-88, shows a significant positive effect (Khan, 1990, Table 2). Pastor (1987) on 1965-81 data obtains a similar result but his tests were confined to Latin American countries. None of the other before-after tests noted earlier yields a significant impact on either BoP indicator and in some cases the sign is 'wrong', *i.e.* the tendency is for the balance to deteriorate.

The ability of programmes to reduce inflation appears even weaker, with results showing the inflation rate to increase as often as it reduces but a virtual absence of significant results.¹⁹ If it is accurate to say that programmes are usually intended to improve on the initial situation and that this is a major reason why governments go to the Fund, and if we confine ourselves to statistically significant results, then it seems that its programmes are often unable to achieve the desired turnaround, although Khan's positive result is the only one which includes much recent information. It should, in any case, be stressed that at least part of the reason for this lies outside the control of the IMF.

A rather closely related issue is the extent to which the programmes achieve their own targets (which, in most cases, seek an improvement on the initial situation). Here again the results are discouraging. They are, however, fairly consistent, with most of the studies listed earlier showing that roughly half of the programme outcomes are below target. The results of the two most recent such studies are typical. Heller *et al.*, 1988 Table 4, shows six out of 12 observations below the current account BoP target and six above; and Edwards, 1989 Table 5, shows 52% on or above target for the same variable.

¹⁷ For a recent such survey see Khan, 1990, especially Table 1, p.208.

¹⁸ It is difficult to judge what is the best way of proceeding on this. As will be shown elsewhere in this study, growth has increased in prominence as an objective of IMF programmes in recent years, particularly in ESAF programmes. At the same time, however, the growth objective remains clearly subordinate to balance of payments goals, even in ESAF programmes. For present purposes, however, it makes little difference whether we classify growth as an objective or a by-product.

¹⁹ A partial exception is in Killick, 1984A, pp.240-42, which on 1970s data found a significant initial reduction but that this was not sustained into the second year.

Neither study tests for the overall balance. Both studies were based on 1983-85 data. The only reported target-actual comparison applied to the overall balance (reported in Reichmann, 1978) produced a similar 50-50 split. The results are similar in the various studies for the inflation rate, although with a tendency for a rather smaller proportion of targets to be met.

A number of lines of explanation are possible here. First, we should recall the observation that quantified targets are only as good as the models which generate them, so the results just described are likely to be influenced by ill-specified models and inaccuracies in the data fed into them. Another explanation which has sometimes been offered is that Fund staff deliberately set ambitious targets in order to influence expectations and thus help to produce an economic improvement. There may be situations in which that would be a reasonable line of action but it could scarcely provide a general reason because to act systematically, or even frequently, in such a way would undermine the credibility of the exercise.

Apart from a natural tendency to over-optimism, a more likely general explanation (confirmed by a number of respondents) is the pressure which Fund missions are often under to reach an agreement in the face of constraints on their ability to change policies, or on the power of those policy changes to bring about improvements. In such circumstances, there is a temptation to massage the figures, to come up with a plausible-seeming set of projected outcomes in support of an agreed set of measures, even though the mission may privately have doubts about the feasibility of the predicted outcomes. Similar forces are at work in heavily-indebted developing countries, where IMF involvement in debt-relief operations is likely to increase pressures to agree programmes even though its missions may privately doubt their feasibility, for there will be much creditor and other pressure to conclude an agreement.²⁰ The setting of unrealistic targets matters, however, because it contributes to the frequency of programme break-down reported in Part II and more generally undermines the credibility of what the Fund is seeking to do.

This brings us to the question whether programme countries do better than non-programme countries. This is what the with-without tests seek to tell us, although we should bear in mind the problems of achieving comparability between the two country samples mentioned earlier, and the consequential bias in the results to exaggerate programme effects.

Here again the results are mixed but are more favourable for programme countries in the more recent tests. Thus, for both the current account and (to a lesser extent) overall BoP indicators Khan (1990) found programme countries to have significantly better outcomes than non-programme countries. Gylfason (1987) similarly obtained a significant positive effect for the overall BoP, on 1977-79 data. Loxley's (1984) tests for low-income developing countries in 1971-82 also showed a better BoP outcome by comparison with

²⁰ This case is argued convincingly by Brown, 1990.

non-programme countries, although these results failed standard significance tests. The results of these and other with-without comparisons are also to the advantage of programme countries when it comes to the inflation record, with every such study showing a lowered inflation rate by comparison with non-programme control groups, albeit with generally low significance levels.

For reasons given at the beginning of this Working Paper, the 'big' question remains the counterfactual: are IMF programmes associated with better BoP and inflation outcomes than would otherwise have occurred? We shall never know, of course, but, aside from in-depth country studies, the most serious attempts to get to grips with this unknowable are the 'comparison-of-simulations' and 'generalised evaluation' studies mentioned earlier. How do these come out?

With mixed results. Taking first Khan and Knight's 1981 and 1985 simulation studies, these find significantly beneficial effects from policy measures of the type usually included in IMF programmes for both BoP indicators and the inflation rate, although this approach does not directly test programme effects *per se*.

The most recent literature has favoured instead the generalised evaluation approach. Goldstein and Montiel's (1986) innovatory use of this technique (on 1974-81 data) unfortunately yielded no statistically significant results for the IMF dummy variable at all. Khan's 1990 study found non-significant improvements in the two BoP indicators in the 12 months following introduction of the programme and significant improvements in the following year. However, he found no significant change in the rate of inflation. The results of a further such test are presented in Part III.

Overall, then, the few available counterfactual tests provide some encouragement for the IMF approach but on a less than overwhelming evidential base.

One of the long-standing areas of controversy about the appropriateness of the Fund's approach to BoP policy in developing-country circumstances is that it is too short-term, with the suggestion that any beneficial effects are ephemeral. What light do the various studies surveyed here throw on the sustainability of programme effects?

Here too the results are mixed. Loxley (1984) examined (with-without) outcomes over one and three years. The overall BoP comparison was more favourable for programme countries in the third year but the opposite was true for the current account and in all cases his results were non-significant. However, he did find a significantly reduced inflation rate in Year 2 - a stronger result than for Year 0.²¹ We have already mentioned that Khan's (1990) generalised evaluation tests obtained rather stronger BoP results in his Year 1 by comparison with Year 0, although this was not the case with inflation. Gylfason's (1987) results indicated a weakening BoP outcome, by both measures, in Year 1 against Year 0. Presumably related to this fact was his further

²¹ Throughout this paper Year 0 refers to the twelve months (or the calendar year in some cases) immediately following adoption of a Fund programme.

finding that, after a substantial drop in Year 0, the expansion of domestic credit had returned to pre-programme rates by Year 1. His finding that the results of programmes which incorporated a currency devaluation were better sustained than those which did not is also pertinent.

Part III also reports generally favourable results on sustainability and, overall, it is fair to say that available evidence tends not to bear out those who claim that the effects of Fund programmes are largely ephemeral.

Another criticism is that the programmes impose excessive socio-economic costs on the countries adopting them. Our examination of this issue starts by summarising the considerable evidence on their impact on economic growth (usually GDP growth) before turning to other pieces of evidence.

The measured impact on GDP growth depends on the type of test. The results of before-after tests of the change in growth are almost unanimous that Fund programmes are not associated with any significant change in the growth rate, positive or negative. The only exception to this was Goldstein and Montiel's (1986) finding of a significantly negative association. Much the same result emerges from with-without comparisons of programme and non-programme countries: no significant differences are observable.

More adverse results emerge from other studies, however. One of the strongest is that the shortfall of actual outcomes against programme targets is the greatest in the case of GDP growth. Heller *et al.* (1988, Table 4) show only five of 13 observations on, or better than, target; Edwards (1989, Table 5) records an average success rate of only 28%;²² Zulu and Nsouli (1985, Table 4) for African countries record growth at or above the target rate in only five of 26 cases. It seems that Fund staff were especially unrealistic in their growth targeting, although we have suggested earlier that growth is not best regarded as a target variable in many Fund programmes, in which case what these findings are recording are more in the nature of systematic forecasting biases.

The work with which Khan is associated does provide grounds for believing there may, indeed, be measurable costs in terms of reduced output. First, the simulation tests undertaken with Knight (1981, 1985) predicted significant short-term reductions in growth as a result of IMF-type demand-management programmes. Second, his 1990 generalised evaluation study found significantly reduced growth rates in both Years 1 and 2 using data for 1973-88 (although not when the data were broken down into sub-periods). Goldstein and Montiel (1986) obtained no significant result on this variable. Special interest lies in the results of a comparison undertaken by Khan and Knight (1985, see Chart 1) between the predicted effects of programmes that were confined to demand-management measures and others which incorporated supply-side measures which would increase investment. Under the latter regime the initial loss of output is smaller, the subsequent recovery steeper and the longer-term growth trend settles down about 2% *p.a.* higher.

²² This and other results reported earlier is a weighted average for 1983-85 calculated from the information in his Table 5.

This result implies the prediction that the Fund's ESAF programmes, which do incorporate greater supply-side conditionality, will be more cost-effective in terms of output growth.

What now of other evidence on programme costs? Here there are only odd bits of evidence. Loxley (1984) tested for association between IMF programmes in low-income developing countries and changes in **saving and investment rates**, finding saving to rise in the first year but to have declined by Year 2, and investment rates to be lower throughout. Although none of these results was statistically significant, they are consistent with Khan's results just reported. They are also consistent with the results of evaluations of the effects of World Bank structural adjustment programmes, which are often linked to IMF programmes and which consistently show reduced post-programme investment levels.²³

A fairly frequent allegation is that Fund-type programmes increase **poverty and/or income inequalities**, but this has proved very difficult to test quantitatively. The most firmly-based evidence bearing upon distributional effects is that provided by Pastor's (1987) with-without tests. His study is particularly concerned to test for correlation between programmes and trends in the share of wages in the functional distribution of income and he finds a strongly significant negative relationship. Although we are not aware of other quantitative studies of this type, his finding is consistent with more qualitative and anecdotal evidence that urban real wages are often adversely affected by IMF-type programmes. However, caution is needed in interpreting this as evidence of adverse net effects of programmes on poverty and inequality. This outcome will depend on the importance, and relative earnings, of the urban wage labour force, and the anatomy of poverty in the country. It is also possible that Pastor's results are stronger because he confined himself to Latin American countries.

The Heller *et al.* (1988) study of the poverty effects of IMF programmes was largely inconclusive, in line with other studies,²⁴ although it did conclude that some programme components had aggravated the plight of certain vulnerable groups. Overall, however, what I wrote on this subject a decade ago still appears generally valid (Killick, 1984A: p.246):

... programme effects are likely to be quite complex. Depending on the characteristics of the economy, the programme in question and the political priorities of the responsible government, the net effect could be to increase or reduce [income] concentration; there is nothing intrinsic to the logic of stabilisation that requires inequalities to increase. Given our findings on other variables, however, it is most likely that the majority of programmes

²³ See Mosley *et al.*, 1991, chapter 9; World Bank, 1989; Corbo and Rojas, 1991; and Corbo and Webb, 1991.

²⁴ See especially the studies sponsored by UNICEF in Cornia *et al.*, 1987 and 1988. These are critical of the effects of specific programmes and of the past neglect of distributional aspects in the design of IMF programmes but are guarded in their general assessments of programme effects.

have no statistically significant effects one way or the other, although there can be specific exceptions in either direction.

Finally, in a study not mentioned hitherto, we have quantitative evidence bearing on complaints that IMF programmes increase **political instability**. There are a number of individual cases where it is well known that attempts to implement the provisions of a Fund programme have led to riots and other destabilising events but are these isolated instances in politically fragile situations, or more characteristic? Once again, there is the difficulty of distinguishing the effects of the programmes from those of the economic crises which often precede the calling-in of the Fund. After attempting to control for this and other non-programme influences Sidell (1988) concludes from a large sample of cases that an IMF programme does not significantly increase or promote political instability. He did not even find any correlation with episodes of collective protest, although he does suggest that governments adopting a single programme (as against repeated ones) are more likely to experience protests. One major limitation of this study should be mentioned, however: it relates only to 1969-77. However, a more up-to-date study is available for 1976-85, but confined to sub-Saharan African countries (Moore and Scarritt, 1990), which similarly (albeit reluctantly) concludes that Fund programmes have no significant impact on the nature of African governments.

Overall, then, the evidence on whether Fund programmes impose heavy costs on the economies affected is fairly weak - which is not perhaps surprising since much of the other evidence already surveyed shows how easy it is to exaggerate the programmes' impact on a rather wide variety of variables. Before we turn for explanations of this relative ineffectiveness, however, there are two other points worth bringing out:

- (a) It can be hypothesised that countries at a lower level of economic development find stabilisation and adjustment particularly difficult because of structural weaknesses and inflexibilities in their economies (Killick *et al.*, 1984A, chapter 8; Killick, 1989, pp.42-44). The most direct evidence on this is that provided by Loxley (1984, see his Table 5.1) who compared results for a sample of least-developed countries with those for a sample used by Donovan (1982), which included a substantial number of middle-income developing countries. For both BoP indicators he found his results were worse than for Donovan's sample, although the results were more mixed for other variables. We might also note the particularly adverse evidence on programme effectiveness provided in the Zulu and Nsouli (1985) before-after and target-actual tests, mentioned earlier, which may well have been contributed to by the restriction of their tests to African (and hence mainly least-developed) countries.²⁵

²⁵ An official listing of (usually publicly available) IMF Working Papers includes an April 1989 paper by Joshua Greene on the effects of Fund programmes in Africa, 1973-86, but the Fund has so far withheld this from public release. It may be conjectured that this is because the study found unacceptably weak programme effects in these countries.

- (b) We may also ask whether there is any discernible trend in programme impact over the last two or three decades. It seems not, although we should bear in mind that only one or two of the studies reported have made much use of data for programmes in the 1980s. The best evidence is provided by Khan (1990, see his Table 6). He disaggregates for 1973-79 and 1980-88, finding a weaker current account but much stronger overall BoP effect in the 1980s. However, only the improvement in current account was significant at the 95% level for the 1980s (although the overall BoP was nearly so), whereas both the current account result and an association with higher inflation were significant for the 1970s. The only other study which encompasses the 1980s is Edwards' (1989) target-actual comparisons but we have already noted that his results are in line with those of other target-actual comparisons, which relate to earlier periods. Once again, we have to report no strong conclusion! A further test of this is reported in Part II.

Given the apparently rather considerable extent to which Fund programmes fail to have their intended results we would like to have evidence on the reasons for the shortfalls. Unfortunately, most of the literature under review here does not address this issue much but there is a certain amount of evidence.

1.4 Sources of difficulty

- *The initial situation and exogenous shocks:*

It is rather self-evident that the severity of the initial situation will have a strong influence on the likelihood of programme 'success'. There is similarly little dispute that large disturbances in the global economy have knocked many programmes off course. This is particularly likely to happen when the borrowing country experiences major and unexpected changes in its terms of trade. Moreover, the effect may be asymmetrical: when there is a large sudden terms of trade deterioration the programme 'fails' because the original policies cannot cope with the now worsened BoP situation. But when the terms of trade suddenly improve programmes are also apt to 'fail', or rather to be abandoned, because a lessening in BoP pressures allows the government to feel it can relax its policies.²⁶

Although the studies examined here provide only limited evidence on the importance of the initial situation and of exogenous shocks, the generalised evaluation of Khan (1990, see his Table 4) does broach these subjects. For each of his target variables he finds that easily the most significant 'explanation' is provided by the previous year's value of the same variable. He also found changes in the terms of trade to have highly significant influences on the two BoP indicators and on GDP growth. Goldstein and

²⁶ For an illustration of this process, relating to Kenya in the 1970s, see Killick, 1984B, chapter 5. See also Killick and Mwegu, 1990.

Montiel's analysis is less explicit on these matters but does show (1986, see their pp. 337-38) how pre-programme conditions can make a major difference to the outcome of before-after and with-without comparisons.

□ *Impact on policies:*

Another possible reason for low programme impact is that they do not have sufficient impact on policy instruments to make much difference to BoP or other outcomes. The first suggestion that this might be the case was given by Reichmann and Stillson's 1978 before-after study. They found there was no significant change in the rate of expansion of total domestic credit (a key policy instrument in IMF programmes) in two-thirds of the cases studied; the same was true for credit to the public sector in half of the programmes, although there was a significant deceleration in two-fifths of the cases. Killick (1984A, Table 7.4) similarly found no significant change in the growth of domestic credit on a before-after basis.

Also on data relating to the 1970s, Gylfason (1987) found a highly significant initial reduction in credit expansion but that by Year 1 this was nearly back to the pre-programme rate, while monetary expansion was a good deal higher (32% *p.a.* against 25% - see his Table 2). In their study of programmes in Africa Zulu and Nsouli (1985, see their Table 4) also provide evidence. In target-actual comparisons they found for total domestic credit that about half of their observations were on target or better, with a similar half-half record in before-after comparisons. The outcome with attempts to reduce budget deficits was much weaker. Only 8/27 actuals were on target or better, and on a before-after basis only 12/29 budget deficits were lower than in the pre-programme year.

We should also mention information provided for programmes in the early-1980s by Heller *et al.* (1988, see their Table 5), which shows 6/11 observations for credit to the public sector at above-target levels. True, their statistics on credit to the private sector reveal only 3/8 above target but that is an ambiguous result because it may well have been a programme objective to raise the proportion of credit going to the private sector. Finally, a study by Cashel-Cordo and Craig (1990) studies the policy leverage of different forms of aid, where IMF credits are included as a form of aid. They found Fund credits to be associated with reductions in both government expenditures and revenues, with a larger relative decline in revenues. This result was statistically significant for stand-by credits but not for Extended Facility credits. Presumably, the decline in revenues was related to slower economic growth. They suggest that Fund credits have substantial leverage in 'buying' quite large fiscal effects but observe that the Fund is better at shrinking the total size of the public sector than at reducing the budget deficit.

□ *Programme execution:*

Overall, then, the result of the above survey of programme impact on policy instruments suggests rather strongly that the Fund is unable to change policies to the extent that it would like to - and to the extent that would be necessary for programme success. One possible line of explanation of this is incomplete implementation of programme provisions. Fund staff members often cite 'lack of political will' as a common reason for programme break-down. Part II demonstrates frequent incomplete implementation of programme provisions, and there is a good deal of other supporting information. The most recent published evidence is provided by Edwards (1989, his Table 4) who examines the compliance record for 34 programmes approved in 1983 in respect of the government budget deficit, total domestic credit and domestic credit to government, finding that, overall, less than half (45%)²⁷ of all such conditions were observed. Observance of programme requirements was particularly weak for the government budget; strongest for credit to government. A similar story is told by earlier studies.²⁸

There is, moreover, evidence that there is a general, though far from perfect, association between implementation and programme results (Killick *et al.*, 1984A, pp.257-59). However, Connors (1979) produced a weaker result. He divided his programme countries into compliers and non-compliers and tested for association between programme results and compliance. Although most of his signs were 'right' only one, relating to inflation, was statistically significant.

□ *Catalytic effects:*

It is a long-standing ambition of the IMF that its programmes should have a catalytic effect, with the 'seal of approval' signified by a programme agreement triggering additional public and private sector capital inflows. A more recent variant on this is that agreement with the Fund is often a formal pre-requisite for debt reschedulings, or debt reductions, through the Paris and London Clubs.²⁹ That programmes do have such catalytic effects is important for their success because the resources of the IMF itself have been allowed to decline to the point that the credits it can make directly available in support of a stabilisation programme are often small in relation to total financing needs.

²⁷ This figure is calculated from data provided in Edwards' Table 4.

²⁸ The results are summarised in Killick *et al.*, 1984A, Table 7.3, p. 253.

²⁹ On this see Rieffel, 1985.

A final hypothesis which we can explore, therefore, is that programmes are relatively ineffectual because they do not bring the intended beneficial effect to the BoP capital account and are thus, in effect, under-funded. However, forming a judgement on this is by no means easy because the capital account includes a variety of types of capital flow and we would not expect them all to be responsive to agreement on a Fund programme to the same extent. There may also be important financial flows 'above the line', most notably interest payments on commercial debt. Indeed, there may be 'perverse' reactions. The clearest case of this was in some heavily indebted countries during the 1980s where IMF credits were, in effect, used to service commercial bank debts.³⁰

Such evidence as there is on this rather underlines the warning that we should not in the general case expect major net catalytic effects. A survey of the evidence relating to the 1970s arrived at this conclusion (Killick *et al.*, 1984A, pp.235-36). Comparison of the results reported above for programme effects on the BoP on the current and overall balances produces a similarly cautionary conclusion. Were there a strong catalytic effect, we would expect the reduction in the overall BoP deficit (or increase in surplus) to be substantially larger than for the current account but, although one or two studies do report such an outcome, no strong such general result emerges from the empirical studies reported here. Moreover, the new results reported in Part III indicate the current account improvement to be rather larger than for the overall balance, suggesting a perverse movement of capital. It may well be, therefore, that programme effectiveness is undermined by continuing shortages of foreign exchange and the depressing effects these are liable to have on export performance, economic activity generally and government revenues.

~~~~~

We now turn from reviewing the evidence in already-published studies to report the results of research undertaken at the Overseas Development Institute, commencing with an analysis of the substantial proportion of Fund programmes which are abandoned before the end of their intended life, and continuing, in Part III, with a before-after analysis of experiences in the 1980s.

---

<sup>30</sup> In his report on the WIDER study, Taylor (1988, p.144) similarly emphasises the difficulties of making any simple statements about catalytic effects. Connections between capital flows and macro stabilisation efforts are complex, he suggests, especially given the turbulence in the 1980s in bank financing and debt servicing.

## II. A STUDY OF PROGRAMME NON-COMPLETIONS

### II.1 Explanation and justification

With patience and ingenuity it is possible to assemble from publicly-available sources a record of the extent to which IMF programmes are drawn down as intended and what follows is based on such information for all programmes approved in 1980-90.<sup>31</sup> The basic premise of this analysis is that the discontinuance of a programme before the end of its intended life is a useful indicator of performance under that programme. Why this should be so requires explanation.

The Letters of Intent which are formally addressed by governments to the Fund management and which set out the agreed policy provisions of the programmes identify a number of performance criteria, which will commonly be monitored on a quarterly basis.<sup>32</sup> Among the most common of these will be ceilings on banking-system credit to the government (or public sector), on total domestic credit, the net domestic assets of the banking system and on various categories of new external borrowings by the government. There may also be specified minimum levels of foreign exchange reserves (a type of BoP target), requirements concerning exchange rate policies and on the government's budget outturn.

A proportion of the credit associated with the programme will be payable at the time of its approval with the balance payable in (not necessarily equal) instalments or tranches. Including the initial instalment, there will typically be about five such tranches. Unless the terms of the agreement are subsequently changed by the granting of waivers or modifications by the Fund, access to outstanding tranches will be conditional on observance of the performance criteria. Thus if, say, credit to government goes above the specified ceiling access to any outstanding tranches of the credit will be suspended (unless a waiver or modification is granted) until the figure is brought back within the ceiling. A high proportion of programmes suspended in this way are subsequently abandoned, although they may be followed quickly by a new agreement.

Interruptions in IMF credit disbursements thus tell us that the government has been unwilling or unable to conform to the agreed terms and that key economic magnitudes have exceeded (or gone below in the case of minima) the levels deemed by the Fund staff to be the maxima consistent with the objectives of the programme. Since most performance criteria refer to policy variables, this further gives us information on the

---

<sup>31</sup> I should like to acknowledge the major contribution of Moazzam Malik to this work. It was he who undertook the detective work on programme break-down and subjected the results to statistical analysis. The chief sources from which the relevant information was gleaned were various issues of the IMF's *Annual Reports*, augmented by issues of the *IMF Survey*. Detailed information on the programmes and their completion status is available from the author on request.

<sup>32</sup> For a more detailed account see Killick *et al.*, 1984A, chapter 6.

impact of programmes on fiscal, monetary and other policy variables. At the same time, our use of the results of this test should be used with caution because it sometimes happens that programmes are allowed to lapse by mutual agreement, chiefly because of unforeseen developments which render the targets or other aspects of the programme unattainable. In these cases programme breakdown does not imply delinquency by either party. On the other hand, while observance of performance criteria may be a necessary condition of programmes achieving their objectives it is certainly not a sufficient condition, so for this reason too our indicator can only be a partial one. Programmes may be mis-specified; policy changes may have insufficient force or be undermined by the adverse effects of other policies; external conditions may deteriorate: for these and other reasons even programmes which successfully meet their performance criteria may often fail to achieve their objectives.

A further factor to bear in mind is that most programmes have in the past been in the form of 'stand-bys' (although that has changed in the most recent years with the adoption of a number of SAF and ESAF programmes) and that the provisions in such programmes typically extend for 12-18 months, at the end of which time all tranches are expected to have been released by the Fund. It ought not, for such programmes, to be an excessively harsh criterion to test whether implementation remains within agreed parameters for such a relatively brief period. The position is different in the cases of EFF and ESAF programmes, which normally have an intended life of about three years. For them we might anticipate a rather higher rate of non-completion.<sup>33</sup>

For the purposes of the present study an uncompleted programme is defined as one in which 20% or more of the total value of the credit remains undrawn. What this means in most cases is that the programme is discontinued at least before the final credit tranche is released. Of course, in many of the cases of non-completion the undrawn balance was considerably greater than 20%.

Although the test just described is a limited one it has the advantage that, once the basic information has been assembled, it is a straightforward matter to look for regularities in the pattern of non-completions. The following analysis is based on information on all the 298 programmes approved in 1980-90. However, as at end-1990 46 of these were still in process of implementation and thus had to be excluded from our calculations. We also excluded ten SAF programmes which had been converted into ESAFs which themselves were still current. It did not seem appropriate to treat such conversions as indicating break-downs. We were thus left with 242 programmes for analysis.

We were, happily, in a position to form an independent judgement on the reliability of our 20% test as an indicator of programme performance. At the time this work was undertaken we were also engaged in the survey of country-specific case studies reported in Working Paper 48. Because these were based on in-depth research by authors

---

<sup>33</sup> The position with SAF programmes is different again, for the level of conditionality attached to these programmes has been rather nominal and there is hence little reason for expecting high break-down rates for them.

knowledgeable about the countries they were analysing, these provide better information on programme performance. This survey took in 17 countries encompassing a total of 48 programmes in the 1980s. We were thus able to compare the information provided in the country studies with the results of our much simpler test based simply on programme completion.

The outcomes matched almost perfectly! Although the work was undertaken separately and the comparison was not made until the rest of the work had been completed, an identical match was obtained on the 29 stand-bys common to both exercises, while there was an only small difference on EFFs, for which the 20% test indicated 13/16 programmes failed during the course of their intended life while the (more accurate) country studies indicated a 15/16 score. There was also a discrepancy in one of the three SAF programmes.<sup>34</sup>

The highly satisfactory outcome of this comparison confirms the accuracy of our information on programme break-downs and supports the usefulness of this as an indicator of programme performance. It thus adds weight to the significance of the results reported above and suggests that this simple test offers a useful low-cost way in the which progress of IMF programmes may be monitored in the future.

What now can we say from an analysis of uncompleted programmes?

## II.2 The results

The results are summarised in Table 2. Just over half (52%) of the 266 programmes were uncompleted in the sense defined. However, the results differed according to programme type, with the best results (48% non-completions) for stand-bys and by far the worst for EFFs (86%). The 38% result for SAFs should probably be discounted because of the small number of observations.

The specially poor result for the EFF programmes can be related to the controversies which surrounded such programmes early in the 1980s. A considerable number of these were signed in the late-1970s and in 1980-81 but thereafter use of this facility was almost abandoned. The IMF staff defended this change of policy by asserting that performance under these programmes had been poor while others (including the present writer) were critical of the change, disputing that the EFF's record was specially weak and deploring a retreat from a facility which permitted the Fund to take at least a medium-term view and to include more supply-side measures in the programmes.<sup>35</sup>

In retrospect it is clear that the staff were right in their judgements about EFF programme success. By our non-completion test, and even though we would expect a higher slippage

---

<sup>34</sup> Details of these comparisons can be provided by the author on request.

<sup>35</sup> On this see Killick *et al.*, 1984A, pp.211-12 and 247-50.

**Table 2: Analysis of IMF programme completion, 1980-90<sup>(a)</sup>**  
(Percentages. Numbers in parentheses are number of programmes)

|               |                                                                                                                                                                                                | <i>Uncompleted (%)</i> |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| 1.            | <b>All programmes (266)</b>                                                                                                                                                                    | <b>52</b>              |
|               | <b>of which:</b>                                                                                                                                                                               |                        |
|               | Stand-bys (220)                                                                                                                                                                                | 48                     |
|               | EFFs (29)                                                                                                                                                                                      | 86                     |
|               | SAFs (16)                                                                                                                                                                                      | 38                     |
| 2.            | <b>By period<sup>(b)</sup></b>                                                                                                                                                                 |                        |
|               | 1980-83 (91)                                                                                                                                                                                   | 44                     |
|               | 1984-86 (68)                                                                                                                                                                                   | 41                     |
|               | 1987-90 (47)                                                                                                                                                                                   | 67                     |
| 3.            | <b>By region<sup>(b)</sup></b>                                                                                                                                                                 |                        |
|               | Sub-Saharan Africa (106)                                                                                                                                                                       | 48                     |
|               | Western hemisphere (47)                                                                                                                                                                        | 52                     |
|               | Asia (26)                                                                                                                                                                                      | 44                     |
| 4.            | <b>By income category<sup>(b,c)</sup></b>                                                                                                                                                      |                        |
|               | Low-income                                                                                                                                                                                     | 46                     |
|               | Lower-middle                                                                                                                                                                                   | 51                     |
|               | Upper-middle                                                                                                                                                                                   | 49                     |
| 5.            | <b>By debt status<sup>(b)</sup></b>                                                                                                                                                            |                        |
|               | Severely indebted (130)                                                                                                                                                                        |                        |
|               | Low-income (75)                                                                                                                                                                                | 51                     |
|               | Middle-income (55)                                                                                                                                                                             | 53                     |
|               | Moderately indebted (45)                                                                                                                                                                       |                        |
|               | Low-income (20)                                                                                                                                                                                | 40                     |
|               | Middle income (25)                                                                                                                                                                             | 52                     |
|               | Others (42)                                                                                                                                                                                    | 38                     |
| 6.            | <b>By dominant export<sup>(b)</sup></b>                                                                                                                                                        |                        |
|               | Primary products:                                                                                                                                                                              |                        |
|               | fuel (11)                                                                                                                                                                                      | 53                     |
|               | non-fuel minerals (38)                                                                                                                                                                         | 41                     |
|               | agricultural (81)                                                                                                                                                                              | 57                     |
|               | Manufactures (25)                                                                                                                                                                              | 44                     |
|               | Diversified exports (28)                                                                                                                                                                       | 38                     |
| <b>Notes:</b> | (a) Analysis excludes SAF programmes converted into ESAF programmes which were still current at the time of analysis. It also excludes other programmes that were still current at April 1991. |                        |
|               | (b) Analysis relates to stand-by programmes only.                                                                                                                                              |                        |
|               | (c) Reported results are the means of results obtained by using 1980 and 1988 country income classifications.                                                                                  |                        |

rate on three-year as against one-year programmes, its record now seems little short of disastrous (a judgement supported by Working Paper 48). It is therefore rather surprising that at the beginning of the 1990s the Fund again began to approve EFF programmes.<sup>36</sup> It will be interesting to see whether they (and ESAF programmes, which are also medium-term) fare better than a decade earlier.

Given the preponderance of stand-bys in the 1980s (220 of our 266 programmes) and the brevity of their intended disbursement period, their 48% non-completion rate is the outstanding fact to emerge from the analysis. The IMF can scarcely be satisfied with a situation in which half such programmes break down so quickly. Moreover, further analysis revealed that 16% of all stand-bys (or 33% of the discontinued ones) broke down almost immediately, with little or no utilisation of the credit beyond the portion that was payable on approval of the programme.<sup>37</sup>

Items 2 to 6 of Table 2 above set out further analyses of our data which, in some respects, enable us to go further than the literature surveyed in Part I. Analysis of these items is confined to stand-bys to avoid the distortions that would be introduced by including EFFs. First, we can ask about completion **trends over time**. From this it emerges as statistically significant (at the 99% confidence level) that non-completion was a specially serious problem in the last years of the period, 1987-90, with a 62% non-completion rate of those programmes approved in this period and not still current, a proportion which rises to an extraordinary 72% if we confine ourselves to the 39 observations for 1988-90.

This is perhaps particularly noteworthy because the period in question was not one of special turbulence in world economic conditions. World output and trade were expanding quite fast, real oil prices were low and, taking the period as a whole, other world commodity markets were less depressed than they had been in the immediately preceding years. Perhaps that was the problem: easier global economic conditions allowed more developing countries to allow their Fund programmes to lapse. This is among our strongest results and shows the difficulties the Fund continues to have in attempting to improve on the poor past record on programme compliance. There is no evidence of an improving trend here; quite the contrary.

We next examined whether programme completion was related to **regional factors**. The hypothesis was that, for different reasons, programmes in African and Latin American countries would have less success than those in Asian countries. In the African cases this could be predicted because of generally deteriorating terms of trade and structurally weak economies; and in Latin America because of the negative effects in many of them of a large debt overhang. Table 2 suggests that there might be something in this, with Asian countries revealing the lowest non-completion rates, but the differences are not large.

---

<sup>36</sup> Three EFF programmes were agreed with developing countries in 1989, of which one (with the Philippines) broke down early in 1991.

<sup>37</sup> The test we applied to ascertain this was that less than 25% of the credit was utilised.

On the hypothesis that the least developed countries would have the greatest difficulties in executing Fund programmes successfully (for reasons sketched earlier) we examined whether programme completion rates were a function of country *per capita* income levels, using the World Bank's system of country income classification. Table 2 shows that this hypothesis was conclusively rejected, with no significant differences in completion rates between the three income categories. A possible reason for this result is that in the 1980s any relationship between programme viability and level of development is liable to have been obscured by the special problems of the heavily-indebted countries, a high proportion of whom are classified in the middle-income categories.

Countries' **debt positions** yield an apparently stronger explanation, however, as can be seen from item 5 of Table 2, even though the results only pass a 90% significance test. While over half of programmes entered into by severely indebted countries broke down, this was true of only 38% of the 'other' group of countries with small external debts. Among the moderately indebted class of countries results appear sensitive to *per capita* income, with middle-income countries recording a much weaker record than low-income countries, although we have no ready explanation for this.

We conjecture that this result is attributable to the already severe strains which the governments of heavily-indebted countries are already under at the inception of the programme. In particular, severe external indebtedness means that large amounts of public resources must be devoted to the servicing of external debt, making it all the more difficult to reduce fiscal deficits in the way that is normally required in IMF programmes. Moreover, the use of currency devaluations is less appropriate in such situations and where the debtor country's amortisation payments have made it a *de facto* net exporter of capital to the rest of the world<sup>38</sup>, and this reduces the number of effective policy instruments available to make a programme work. Finally, the economic hardships already imposed on the populace by the debt burden are likely to make strict adherence to the further austerities of IMF-sponsored programmes politically very difficult.

One other point of note is the heavy concentration of stand-bys on severely-indebted countries in the 1980s, shown in the table: 130 out of the 217 classified by debtor status. Since many of these programmes were necessary preconditions for debt re-scheduling agreements in the Paris and, to a lesser extent, London clubs, the results obtained here suggest that in many cases the debt relief secured was not enough to be consistent with programme success.<sup>39</sup> These results are further consistent with the generally weak findings on the catalytic effect of Fund programmes reported in Part I and later.

Rather closely related to the above is the question of the **adequacy of the IMF's own credits**. It is sometimes suggested that these are small in relation to need, especially when one takes into account return flows to the Fund in respect of earlier credits. We

<sup>38</sup> On this point see Reisen and van Trotsenburg, 1988.

<sup>39</sup> For a detailed treatment of this theme, with reference to African countries, see Martin (forthcoming).

therefore undertook an analysis of the magnitude of the credits drawn down in our sample of 38 programmes and related these to BoP magnitudes. The results are summarised in Table 3.<sup>40</sup>

| <b>Table 3: Indicators of the adequacy of Fund credits, 1979-89</b><br>(percentages) |                                                    |                               |                     |                |                     |
|--------------------------------------------------------------------------------------|----------------------------------------------------|-------------------------------|---------------------|----------------|---------------------|
|                                                                                      | <u>Annualised mean use of IMF credits as % of:</u> |                               |                     |                |                     |
|                                                                                      | <i>GDP</i>                                         | <i>Base-period values of:</i> |                     |                | <i>Return flows</i> |
|                                                                                      |                                                    | <i>current</i>                | <i>overall</i>      | <i>imports</i> | <i>during</i>       |
|                                                                                      | (1)                                                | <i>a/c deficit</i>            | <i>BoP deficit</i>  | (4)            | <i>programme</i>    |
|                                                                                      |                                                    | (2)                           | (3)                 |                | (5)                 |
| All programmes                                                                       | 1.8                                                | 29.3                          | 22.6                | 9.9            | 1205                |
| Completed programmes                                                                 | 2.6                                                | 48.3 <sup>(a)</sup>           | 73.4                | 16.9           | 1105                |
| Uncompleted programmes                                                               | 1.3                                                | 18.6 <sup>(a)</sup>           | -6.2 <sup>(b)</sup> | 6.0            | 1261                |
| 1979-82 programmes                                                                   | 2.1                                                | 25.7                          | 0.2                 | 7.9            | 862                 |
| 1983-85 programmes                                                                   | 1.4                                                | 33.4                          | 47.5                | 12.2           | 1588                |

*Notes:* (a) The difference between this pair of average values is significant at the 90% level.  
(b) The negative sign indicates that there was an initial overall surplus.

Of course, what constitutes 'adequacy' is a matter of judgement and will vary from case to case. Nonetheless, if we take all the programmes together it can be seen that credits utilised were equivalent to three-tenths of the pre-existing balance of current account, nearly a quarter of the overall BoP and nearly 10% of total imports. They were therefore far from being insignificant. Moreover, the credits utilised exceeded by a factor of 12 the return flow of amortisation ('repurchase') payments in respect of previous programmes (column (5)). The Fund was thus providing quite substantial assistance on a net basis. This needs to be qualified in two respects, however. First, the adequacy indicators were substantially lower in 1979-82 than in the following years. Second, but not shown in the table, the net flow was substantially smaller for countries which had just successfully completed an earlier programme. For them the net flow was only 231% of repurchases.

<sup>40</sup> The values used in the compilation of Table 3 are annualised averages of actual uses ('purchases' in the Fund's parlance) of credits during periods when the programmes in question were operative. In cases where programmes were suspended or abandoned, only purchases during the period when the programmes were active have been included and the period adjusted accordingly when annualising the data. The base value is taken as the mean annualised value of the variable in question during the two years immediately preceding the programme.

Of more direct relevance to the present discussion, however, are the differences between the positions of the countries which did and did not complete their programmes in the period under examination, for it can be seen that most of the indicators of adequacy are substantially smaller for the non-completers, even though these annualised data are adjusted for the effects of programme suspensions and cancellations. In other words, relative to imports and BoP balances, the non-completers received substantially smaller credits than the others - a difference that was, however, only significant at the 90% level. Although it can be no more than suggestive, this evidence is consistent with the hypothesis that inadequate supporting finance is an important reason for programme break-down.

Finally, we looked for any association between frequency of programme break-down and **type of export** (item 6). The hypothesis here was that exporters of manufactures would have relatively high completion rates because, being sellers on generally buoyant world markets, they would find it easier to boost export earnings by comparison with exporters of (non-fuel) primary products. This hypothesis receives little support, however. Exporters of manufactures turn out to have only moderately below-average break-down rates, while exporters of non-fuel mineral products have a rather better record. The record of agricultural exporters does conform to our expectations, however, with a 57% break-down rate (statistically significant at the 95% level). Interestingly, the lowest rate belonged to countries classified (by the IMF) as having diversified export bases.

There is a further use to which we can put the information assembled for the above analysis. It pertains to the **sustainability of programmes**. We should recall here that IMF assistance is intended to provide temporary support to governments seeking to restore viability to their countries' BoP. It was not expected to be providing frequent credits over a sustained period. From our inventory of 1980-90 programmes, however, it is clear that repeated assistance is quite frequent. In fact, over these 11 years no less than nineteen countries had six or more programmes approved by the Fund, encompassing 131 programmes, or 44% of the total for the period. Such programmes were often approved back-to-back or with only brief intervening periods. The record was nine programmes in the eleven years, a distinction shared by Madagascar, Senegal and Togo - all from the African region, be it noted, and two from the Franc zone group of countries. While the Fund should be given credit for its persistence with these countries, its founding fathers cannot be happy that its programmes were apparently unable to bring any sustained improvements to their payments positions.

### III. A BEFORE-AFTER ANALYSIS OF THE 1980s

#### III.1 Nature of the study

We turn next to offer some additional new empirical information concerning programmes adopted in the 1980s. In the course of conducting the research on which Working Paper 48 is based, statistical information was collected on a range of macroeconomic variables for each of the countries studied and in the pages that follow we utilise this information to present additional quantitative information, based on before-after tests. The limitations of such tests have already been described and should be borne in mind in what follows. However, we have also attributed more usefulness to them than is currently fashionable. Details of the data sources and tests employed are set out in the Appendix.

What follows has some advantages over most of the comparable literature. It is more up-to-date, being based on programmes commenced and completed (or abandoned) in 1979-85. It examines a wider range of BoP variables so as to facilitate a more rounded view of programme effects. It includes an analysis of changes in domestic absorption and its chief components, on the grounds that a strengthening of the BoP is likely to require a reduction in total absorption relative to GDP and in order to examine further where cuts in absorption most make themselves felt. It takes us further into the neglected area of programme effects on key policy variables. It studies the effects of programmes over a longer period (up to four years from commencement of the programme) in order to better test the sustainability of programme effects. It tests for differences in results between programmes that are, or are not, completed. Finally it also tests for whether the fact that there was another IMF programme immediately prior to the current programme had any significant effect.

#### III.2 The results

The results are summarised in Table 4. The two right-hand columns of this differentiate the Year 1+2 results according to whether the programmes in question were completed or not, using the same 80% drawn-down cut-off as that utilised in Part II above.

##### □ *Balance of payments effects:*

Improvements to the BoP were among the strongest of all our results, with appreciable and statistically significant improvements in both the overall and current account balances. The effects were relatively weak for the programme year (Year 0) but were strong for the following three years. A quantitatively larger reduction in the current account deficit was achieved in countries which completed their programmes but, curiously, the opposite was true with respect to the overall BoP.

We should note, moreover, that the BoP results were obtained in the face of rather strongly adverse movements in the commodity terms of trade (line 4), which by Year 3 had, on average, deteriorated by a full 8%. Observe here that the experiences under the uncompleted programmes was substantially worse than with completed programmes, which suggests strongly that worsening trade conditions were a source of programme breakdown.

One possibility is that the current account results were secured simply on the basis of an import squeeze (in which case we would have to be careful about calling this an improvement) and line 6 of the table provides information on this. The statistics there show that there was a significant tendency for import volumes to be cut in the programme year, a tendency (not shown in the table) concentrated in the most recent years.<sup>41</sup> The impact was not large, however, nor was the result statistically strong except in the programme year. A possible explanation for this is that the crisis conditions in which many programmes are introduced resulted in reduced import volumes in the year or two immediately preceding the programme, because of reduced credit-worthiness, depleted reserves, growing prior claims of debt-servicing obligations, *etc.* We therefore examined the behaviour of imports in the years immediately preceding programme adoption. This revealed that there was a modest tendency for imports to be reduced in the pre-programme year but it was not statistically significant. All in all, the worst fears that IMF programmes lead to sustained import strangulation were not borne out.

On a more positive note, we found rather strong evidence that programmes are associated with improved export performance, as indicated by growth in export volumes (see line 1 of the table). Given normal supply response lags, it is not to be expected that there would be any large immediate export response but we found consistent trends (significant at the 99% level or above) from Years 1 and 2 through to Year 4, particularly among those countries which had previously completed programmes.

It is commonly claimed that one of the chief values to a country of agreeing a programme with the IMF is that this has a catalytic effect on net capital inflows. Indeed, it is important that it should do so, given the limits on the amount of supporting finance the Fund can itself lend. Items 3 and 4 of Table 4 provide evidence on this (treating official transfers as capital transactions). As can be seen, the changes recorded are quantitatively quite

---

<sup>41</sup> For programmes commenced in 1983-85 there was a mean reduction in import volumes of nearly 9% on the average values in the immediately following two years. Using a one-tailed test, this was significant at the 95% level. It should be added, however, that there was a considerable spread of values around the means.

**Table 4: Results of before-after tests on sixteen developing countries with IMF programmes commenced in 1979-85<sup>(1)</sup>**

| <i>Variable</i>                                            | <i>Base value</i> |                     | <i>Differences from base value</i> |                      |                                       |                                         |
|------------------------------------------------------------|-------------------|---------------------|------------------------------------|----------------------|---------------------------------------|-----------------------------------------|
|                                                            | <i>(Yr -1,-2)</i> | <i>Yr 0</i>         | <i>Yrs 1,2</i>                     | <i>Yr 3</i>          | <i>Completed programmes (Yrs 1,2)</i> | <i>Uncompleted programmes (Yrs 1,2)</i> |
| <b><u>Balance of payments indicators</u></b>               |                   |                     |                                    |                      |                                       |                                         |
| 1. Overall BoP (as % GDP, change in reserves)              | -1.2              | +0.4 <sup>d</sup>   | +1.2 <sup>a,d</sup>                | +1.9 <sup>a,d</sup>  | +0.7 <sup>d</sup>                     | +1.5 <sup>a,d</sup>                     |
| 2. Current a/c (as % GDP, exc. official transfers)         | -11.2             | +1.1                | +3.0 <sup>a,d</sup>                | +3.8 <sup>a,d</sup>  | +4.4 <sup>b</sup>                     | +2.2 <sup>b,d</sup>                     |
| 3. Official transfers (as % GDP)                           | 2.8               | -0.2                | +0.0 <sup>d</sup>                  | +0.0 <sup>d</sup>    | -1.4 <sup>d</sup>                     | +0.8 <sup>d</sup>                       |
| 4. Capital account                                         |                   |                     |                                    |                      |                                       |                                         |
| (a) direct foreign investment (as % GDP)                   | 0.4               | +0.0                | +0.1                               | +0.3 <sup>b</sup>    | -0.0                                  | +0.2                                    |
| (b) Net long-term loans (as % GDP)                         | 5.5               | -0.0                | -1.2 <sup>d</sup>                  | -2.0 <sup>a,d</sup>  | -1.1 <sup>d</sup>                     | -1.3 <sup>d</sup>                       |
| 5. Terms of Trade index (% change) <sup>(2)</sup>          |                   | -3.8 <sup>a,d</sup> | -6.5 <sup>a,d</sup>                | -8.2 <sup>a,d</sup>  | -2.5                                  | -8.8 <sup>d</sup>                       |
| 6. Import volume index (% change)                          |                   | -4.6 <sup>b</sup>   | -3.1                               | -4.1                 | +0.2                                  | -5.0                                    |
| 7. Export volume index (% change)                          |                   | +3.5 <sup>d</sup>   | +9.6 <sup>a,d</sup>                | +11.7 <sup>a,d</sup> | +8.7                                  | +10.2 <sup>a,d</sup>                    |
| <b><u>Inflation and growth</u></b>                         |                   |                     |                                    |                      |                                       |                                         |
| 8. Increase in consumer prices (% p.a.) <sup>(3)</sup>     | 24.8              | +1.7                | -1.1 <sup>c</sup>                  | -0.4                 | -12.7 <sup>e</sup>                    | +5.6                                    |
| 9. GDP growth (% p.a., at constant prices)                 | 2.1               | -0.0                | +0.6                               | +1.2                 | +1.8                                  | -0.0                                    |
| <b><u>Domestic absorption</u></b>                          |                   |                     |                                    |                      |                                       |                                         |
| 10. Total absorption (as % GDP)                            | 107.5             | -2.5 <sup>a,d</sup> | -3.4 <sup>a,d</sup>                | -3.3 <sup>a,d</sup>  | -1.2                                  | -4.6 <sup>a,d</sup>                     |
| 11. Private consumption (as % GDP)                         | 71.6              | -0.3                | -0.4                               | +0.2                 | +1.5                                  | -1.5                                    |
| 12. Government consumption (as % GDP)                      | 13.7              | -0.4                | -0.3                               | -0.4                 | -1.0 <sup>b</sup>                     | +0.1                                    |
| 13. Fixed investment (as % GDP)                            | 21.2              | -1.9 <sup>a,d</sup> | -3.1 <sup>a,d</sup>                | -3.8 <sup>a,d</sup>  | -3.4                                  | -2.9 <sup>a,d</sup>                     |
| <b><u>Policy variables</u></b>                             |                   |                     |                                    |                      |                                       |                                         |
| 14. Real effective exchange rate index (% change)          |                   | -7.1 <sup>b,d</sup> | -11.0 <sup>a,d</sup>               | -15.0 <sup>a,d</sup> | -18.1 <sup>b</sup>                    | -7.0 <sup>d</sup>                       |
| 15. Total domestic credit                                  |                   |                     |                                    |                      |                                       |                                         |
| (a) growth rate (% p.a.)                                   | 25.8              | -3.4                | -3.3 <sup>d</sup>                  | -4.2 <sup>b,d</sup>  | -3.3                                  | -3.2                                    |
| (b) as % GDP                                               | 42.4              | -0.4                | -0.7                               | -1.8                 | +2.0                                  | -2.4                                    |
| 16. Private sector credit                                  |                   |                     |                                    |                      |                                       |                                         |
| (a) growth rate (% p.a.)                                   | 23.1              | +0.4                | +0.8                               | +0.8                 | -3.6 <sup>d</sup>                     | +3.2                                    |
| (b) as % GDP                                               | 18.9              | -0.8                | -0.8                               | -0.2                 | +0.8                                  | -1.6                                    |
| 17. Credit to central government (as % GDP) <sup>(4)</sup> | 18.2              | +0.2                | +0.0                               | -1.5                 | +1.0 <sup>d</sup>                     | -0.6                                    |
| 18. Central Government budget deficit (as % GDP)           | -6.7              | +0.2                | +1.4 <sup>a</sup>                  | +1.6 <sup>a</sup>    | ...                                   | ...                                     |

*Notes:* (a) Significantly different from zero, under a one-tailed t-test, at the 99% confidence level.  
(b) Significantly different from zero, under a one-tailed t-test, at the 95% confidence level.  
(c) Significantly different from zero, under a two-tailed t-test, at the 95% confidence level.  
For line 8, significance tests are based on log-linear data.  
(d) Significant proportion of programmes with either a positive or negative change, as compared to a hypothesised equal proportion.  
(1) See appendix for notes on the data and details of the various tests.  
(2) Completed and uncompleted programme values are significantly different from each other at the 90% confidence level.  
(3) As for (2), but at the 95% confidence level.  
(4) Year-to-year changes in credit to government were too variable for it to be meaningful to calculate changes in growth rates.

small, although several are statistically significant.<sup>42</sup> The net effect is a reduced capital inflow. One possibly important explanation of this finding is the much smaller initial deficit on the overall account - only 1.2% of GDP. However, the main explanatory factor would seem to be the increase in net repayments of long-term loans - *i.e.* a substantial part of the improvement in the current account was used to finance the repayment of foreign loans and was not rewarded by increases in disbursements (although this may be a much longer-term result). This result is particularly surprising given the frequent linkage of IMF programmes to debt rescheduling agreements (which would reduce the level of principal repayments). In effect, Fund credits were often used to repay other creditors. We should bear in mind, however, that we are not able to make comparison with non-programme countries, which may have experienced even higher rates of net loan repayments.<sup>43</sup>

□ *Inflation and growth:*

Weak results are also revealed with respect to inflation (line 8), with only the smallest reductions on the pre-programme levels. Over 40% of the programmes were associated with an increase in the inflation rate - even in the longer term.

The longer-term results for real GDP growth are more interesting as they suggest that after three or four years programmes may be associated with increases in growth rates. This runs contrary to the findings reported in Part I, where the general conclusion was that the short-run effects were negative. However, the results reported here only offer weak evidence - the increase after four years is only significant at 90% confidence levels. There is also some weak evidence that completed programmes are associated with longer-term improvements in GDP growth rates (statistically significant in Year 3 at the 90% confidence level).

---

<sup>42</sup> One difficulty here relates to poor data and the variety of definitions for the overall balance. The results presented here refer to the change in the official foreign reserves. Using the IFS performance balance definition yields an even smaller and insignificant improvement.

<sup>43</sup> The apparently statistically significant doubling of direct foreign investment unfortunately only illustrates the dangers of relying on t-statistics when working with a small sample. If the two programmes agreed with Dominica are excluded the increase becomes negligible. A similar explanation accounts for the bizarre result that official transfers decrease for completed programmes but increase for uncompleted programmes. This result is entirely the result of large changes occurring just in Dominica. Excluding Dominica reveals that for the rest of the sample there was a significant increase in official transfers, increasing by 30% on average four years after an agreement and increasing to some extent for three quarters of all programmes.

The GDP results are surprising, particularly in view of the reduced levels of fixed investment reported below, and imply that programmes have some beneficial impact on the efficiency of resource utilisation.

□ *Domestic absorption:*

In the absence of any strong catalytic effect, absorption has to be reduced relative to GDP if the BoP is to be strengthened. That programmes are associated with such a reduction is among our strongest results (line 10), with substantial and significant changes over the entire post-programme period. If cause-and-effect are at work here, this must be rated a programme success - but note the odd result that the largest reduction is in respect of uncompleted programmes, with only a non-significant fall in the case of completed programmes. This could be because initial domestic absorption as a proportion of GDP was on average four percentage points higher in uncompleted programme cases than for completed programmes, which reinforces the view that programme success is related to the severity of the initial situation.

But on which component of absorption do the reductions chiefly fall? The answer is that the brunt falls on fixed investment, which declines substantially and significantly over the whole period. Overall, the programmes appear unable to exert any appreciable squeeze on private or public consumption, although there is a shift in its composition in favour of the private consumer in the case of completed programmes. The adverse impact on investment is consistent with results reported in Part I and in evaluations of World Bank programmes. Where programmes have the effect of bringing excess capacity into utilisation, temporary reductions in investment can be consistent with continued economic growth but if investment remains depressed this becomes a serious source of difficulty for policy makers. The results in Table 4 suggest that investment not only remains depressed but actually continues to decline.

□ *Impact on policy variables:*

It was also suggested earlier that their relative lack of impact on key policy instruments provided an explanation why programmes often do not meet their objectives. Items 14-18 in Table 4 provide evidence on this. The strongest result is with respect to the real exchange rate (line 14), which is shown as being depreciated by an average of 7% in the programme year, a depreciation which is not merely sustained but actually deepened during the following three years, and which is particularly large for programmes which run their intended term. This no doubt helps to explain the substantial improvement in export performance reported earlier.

These results, moreover, have generally high levels of significance. Although we cannot make any direct comparison with non-programme countries in this period, it is worth noting that our average rate of -11% *p.a.* is significantly larger than the average real devaluation rate of -1.2% *p.a.* recently reported for a much larger sample of developing countries (84 in total) over the same period (Lynn and McCarthy, 1989) - a rate of depreciation influenced by inclusion of countries featuring also in our sample.

Given the importance of the exchange rate as a policy variable assisting both stabilisation and structural adjustment, the Fund's influence on this instrument is an important finding.

The impact on domestic credit, which is a central feature of IMF programme design, will be regarded by it as less satisfactory. While there is some reduction in the rate of credit expansion and the value of credit relative to GDP, the effects are small and non-significant. Contrary both to expectations and Fund intentions, there is no significant reduction in the share of total credit going to the central government *vis-à-vis* the private sector. More consistent with programme objectives is the significant reduction in budget deficit achieved in Years 1+2 and 3 (although the number of observations in this case is small).

□ *Sustainability:*

One of the ways in which the tests reported here differ from most of the existing literature is in permitting an examination of programme effects over a longer period, although the small size of our sample should be borne in mind here. It is quite common for such tests in the past to have been confined to the programme year (Year 0 in our terminology) and perhaps the year after. Comparing the Year 0 with the following two columns it is evident that tests which are confined to that year are liable to under-state the extent to which IMF programmes are associated with changes in the pre-programme situation, for our results for Years 1+2 and even Year 3 are generally larger and more significant than for Year 0.

In fact, the results for Year 0 are generally weak. If we are willing to adopt the language of cause-and-effect, then it appears from these results that programmes were in the 1980s only able to bring quick changes to the exchange rate (devaluations were often 'prior actions', *i.e.* had to be undertaken before the IMF would approve a programme) and investment (line 13). A substantially wider range of significant results is obtained for the later years, which is perhaps surprising, given that the average programme life in our sample, after allowing for cancellations, was only 18 months. The extent to which BoP improvements were sustained into Year 3 is particularly noteworthy.

□ *Effects of non-implementation:*

There is a strong presumption that the extent of slippage in policy implementation is greater in the case of uncompleted programmes, for reasons given earlier, although some of the policies will have been implemented in virtually all cases, since it is common for the Fund to require a government to undertake some policy measures before a programme will be approved (so-called pre-conditions or prior actions). How much difference does non-completion make?

First and curiously, it seems that, relative to GDP, credit is restrained more in the non-completion cases, overall and with respect to both major components of credit (although the results are statistically non-significant), which raises the questions why the programmes were suspended and whether it was because they incorporated particularly severe, and unattainable, credit ceilings. More in line with expectations, programme non-completion is associated with substantially smaller real exchange rate depreciations, which suggests that government reluctance to devalue may be a reason for break-downs, although much of the statistical difference between these sub-groups was the result of Ghana's huge devaluation. Unfortunately, there were insufficient observations on budgetary effects for the results to be disaggregated between completed and non-completed programmes.

Turning from policy instruments to target variables, while the BoP current account result is smaller in the case of uncompleted programmes, such a result is predictable from the worse terms of trade experiences of the non-completing countries. As might therefore be predicted, the import squeeze is concentrated on the non-completers; their net increase in capital outflows is smaller, and the overall BoP result is actually larger and more significant than with completed programmes.

Turning to the domestic economy, the inflation record is a good deal worse in the case of uncompleted programmes, with a substantial (though non-significant) rise against a handsome reduction in the other group. The completers also have a (non-significantly) larger increase in GDP growth. Finally, and puzzlingly, the reduction in total absorption is particularly large and significant in the non-completion cases, with the brunt again mainly falling on investment.

All in all, then, if we focus on the BoP record as the main focus of the IMF it is not obvious that governments who complete their programmes can expect to get superior results. We tested for the statistical significance of all the pairs of observations in the 'completed programmes' and 'uncompleted programmes' columns but only the line 5 results on the terms of trade were significantly different, even at a 90% confidence level. That

non-implementation should not have more clearly adverse consequences rather under-scores the limited association between IMF programmes and improvements in economic indicators.

□ *Effects of past relationships:*

Finally, although not reported in Table 4, we examined whether programme outcomes were correlated with past recourse to the IMF. For this purpose we classified cases into those which were not preceded by a programme, those that were preceded by an uncompleted programme, and those preceded by a fully drawn-down programme. The broad hypothesis was that countries which had previously worked successfully with the Fund were likely to have the least severe programmes and/or to get the most favourable results. There were significant differences among these three categories for a number of our variables, of which the most noteworthy were:

- The countries which had previously completed a programme achieved the largest reductions in their current account and overall BoP deficits. They (and to a lesser extent the countries which had previously had a programme but had not completed it) also achieved substantially greater growth in export volumes, which points up the time lag which is liable to exist between undertaking adjustment measures and stimulating improved export performance.
- The countries which had not previously had an IMF programme incurred the largest reductions in total absorption (especially in Years 0 to 2) and in investment. They also suffered the severest short-term (Year 0) squeeze on credit and on import volumes, although this was not sustained in later years.

~~~~~

The reader doubtless feels in need of a summing-up at this point. However, Working Paper 48 presents further materials, based on in-depth case studies rather than quantitative analyses, which provide further information on the effects of Fund programmes. We therefore defer a summing up until the conclusion of that paper. However, to respond to the question in the title of this paper, it is already clear that we can know quite a lot, provided we are willing to be satisfied with reasonable levels of probability rather than certainty. It is also clear that the evidence is mixed, does not provide exclusive support to any single school of thought, and that it is easy to exaggerate the amount of impact which the IMF's programmes have, for good or ill.

METHODOLOGICAL APPENDIX

Sample selection and data

The before-after analysis is based on 38 programmes (observations), in 16 developing countries, that commenced during 1979-85. The countries were:

| | | |
|---------------|----------|--------------------|
| Bangladesh | Brazil | Costa Rica |
| Côte d'Ivoire | Dominica | Dominican Republic |
| Ghana | Jamaica | Malawi |
| Mexico | Morocco | Pakistan |
| Philippines | Somalia | Sudan |
| Tanzania | | |

The companion paper also includes materials on The Gambia. This country is not covered in the statistical analyses, however, due to lack of data. As is explained in the companion paper, the country sample cannot be described as random as selection was substantially influenced by the availability and quality of literature on countries with IMF programmes. Attempts were made to avoid bias but these were not fully effective: relative to the total population, the sample contains too many EFF agreements, which results in too high a programme breakdown rate - 63% in our sample against 50% for all 242 agreements over the 1979-89 period.¹

The two principal data sources were the *International Financial Statistics* of the IMF and the World Bank *World Tables*, though these were supplemented by other sources as necessary. Care was taken to ensure comparability within each of the time series. Our tests required data covering at least two years before and four years after the inception of the programme. However, this was not always available. In the case of incomplete data, the number of observations was reduced to below 38. For example, the results reported for the 'Central Government Budget Deficit' (line 18, table 4) are limited to 21 programmes (observations) in 9 countries due to the lack of adequate data. Furthermore, in a number of cases, lack of data left us with fewer observations in particular periods of interest. For example, owing to the absence of 1989 data for Import and Export Volumes, the 36 available observations are further reduced to 31 for Year 3 results. There were also difficulties with some of the data series, particularly with respect to the overall BoP and the government budgetary balance, which were either incomplete or for which different (official) sources provided markedly different statistics.

The restrictions imposed by the availability of data and other materials means that our work is likely to suffer from some sample selection bias. Unfortunately, it has not been possible to test for this, beyond the rudimentary check mentioned above. We have thus been cautious in generalising our results for countries outside the sample described above.

¹ Further details about the sample selection and data base are provided in the companion paper.

The before-after method

Annual data were adjusted to allow for the month in which the programme started. To give a concrete example, in March 1985 Costa Rica began a stand-by agreement due to run until March 1986. Thus, the base values were calculated as the annual average over April 1983-March 1985 (the two years immediately preceding the programme), the Year 0 outcome over April 1985-March 1986, the Years 1 & 2 outcome as the annual average over April 1986-March 1988, and finally, Year 3 over April 1988-March 1989. Calculated this way the base value for Costa Rica's Overall BoP was a surplus of 1.2% of GDP, as compared with 2.7% for the first twelve months of the programme (Year 0), an annual average of 1.9% for the next two years (Years 1 & 2), and 4.0% for the fourth year (Year 3). Alternative bases and periods of analysis were tried, but the results did not prove very sensitive to variations in approach.

We then calculated the difference between these outcomes and the base for each observation, and aggregated so as to calculate an overall mean. Other averages were also calculated corresponding to whether the programme was fully completed or not, whether the programme was preceded by another IMF agreement (and if so, whether it was completed or not), and whether the agreement commenced in the 1979-82 period or 1983-85. This was done for each of the variables under investigation.

Statistical tests

The results were then subjected to a range of tests. First, we tested for the significance of the difference between Year 0, Years 1 & 2, and Year 3 outcomes and the base value - in other words, whether the difference between the outcomes and the base was significantly different from zero. For various data series, we had strong *a priori* reasons to expect a change in a particular direction, and hence one-tailed t-tests were applied to these series. An example is the Current Account Deficit (line 2, table 4) - both theoretical reasons associated with an IMF package and previous evidence predict an improvement in the deficit.² By contrast, we had no *a priori* reasons to expect the Terms of Trade Index (line 5) to change in a particular direction, and hence a two-tailed t-test was used to evaluate whether the change was significantly different from zero or not.

Second, we tested for the significance of the difference between the proportion of programme countries experiencing an increase (or reduction) in each variable and the null hypothesis of a presumed random outcome of 50% showing improvements and 50% showing deteriorations. (The presumption of a 50:50 split could be considered a rather severe test given that IMF programmes are often agreed in crisis conditions; an alternative presumption would be 100% of programme countries showing a deterioration). The test was based on the exact binomial distribution (and not the normal approximation to the binomial) adjusted for different sample sizes.

² Other variables tested under a one-tailed hypothesis were lines 1, 4(a), 6, 7, 10-13, 14, 15(a), 17(a), and 18. All other variables were tested under a two-tailed hypothesis.

Third, the differences between the Year 1 & 2 outcomes of completed and uncompleted agreements were tested for significance. A standard t-test was used -this, however, is valid only if the two series have a common variance, and hence, this latter hypothesis was tested prior to the t-test.³ Only two of the data series (lines 5 and 8) exhibited such significance. Indeed, the absence of significance in other variables would suggest that full or partial implementation has little substantive effect on the outcome of target variables.

Other test statistics were also calculated, to examine whether the variables were normally distributed (Jaques-Bera), whether there was a significant change in the variance of the sample before/after (Goldfeld-Quandt) and whether there was a significant difference in the means (Chow II). These tests, however, are not reported as the results were generally inconclusive. Only in the case of inflation rates, which have a distinctly non-normal distribution, did these tests prompt an investigation of alternative functional forms. Log transformation produced a more 'orderly' data set, and the significance tests reported in Table 4 are on the basis of this transformed data - results from the untransformed data have the same signs but lower significance. Such an arbitrary procedure is, of course, open to objection.

Further details on the data, sources and tests can be provided on request.

³ Normality was assumed throughout, though significant results were then tested to confirm the assumption before reporting.

REFERENCES

- Brown, Richard (1990); 'The IMF and Paris Club rescheduling: a conflicting role?', *Working Paper No. 30*, The Hague: Institute of Social Studies; May.
- Cashel-Cordo, Peter and Craig, Steven G. (1990); 'The public sector impact of international resource transfers', *Journal of Development Economics*, 32.
- Connors, Thomas A. (1979); 'The apparent effects of recent IMF stabilization programs'. Washington: Federal Reserve *International Finance Discussion Paper No. 135* (mimeo); April.
- Corbo, Vittorio and Rojas, Patricio (1991); 'World Bank supported adjustment programs: country performance and effectiveness', World Bank *PRE Working Paper No. WPS 623*; March.
- Corbo, Vittorio and Webb, Steven B. (1991); 'Adjustment lending and the restoration of sustainable growth', *Journal of International Development*, 3(2); April.
- Cornia, G.A.; Jolly, R.; and Stewart, Frances (eds) (1987); *Adjustment with a Human Face: Protecting the Vulnerable and Promoting Growth*. Oxford: Clarendon Press.
- Cornia, G.A.; Jolly, R.; and Stewart, Frances (eds) (1988); *Adjustment with a Human Face: Ten Country Case Studies*. Oxford: Clarendon Press and UNICEF.
- Donovan, Donal J. (1982); 'Macroeconomic performance and adjustment under Fund-supported programmes: the experience of the seventies', *IMF Staff Papers*, 29(2); June.
- Edwards, Sebastian (1989); 'The IMF and the developing countries: a critical evaluation', Carnegie-Rochester Conference Series on Public Policy No. 31. North-Holland.
- Goldstein, Morris and Montiel, Peter (1986); 'Evaluating Fund stabilisation programs with multicountry data: some methodological pitfalls', *IMF Staff Papers*, 33(2); June.
- Goreux, Louis M. (1989); 'The Fund and the low-income countries', in Gwin and Feinberg *op cit*.
- Gwin, Catherine and Feinberg, Richard (eds) (1989); *The International Monetary Fund in a Multipolar World: Pulling Together*. Washington: Overseas Development Council.

- Gylfason, Thorvaldur (1987); 'Credit policy and economic activity in developing countries with IMF stabilisation programs', *Princeton Studies in International Finance No. 60*. Princeton University; August.
- Heller, Peter S.; Bovenberg, A. Lans; Catsambas, Thanos; Chu, Ke-Young and Shome, Parthasarami (1988); 'The implications of Fund-supported adjustment programs for poverty', *IMF Occasional Paper No. 58*. Washington: IMF; May.
- Khan, Mohsin K. (1990); 'The macroeconomic effects of Fund-supported adjustment programs', *IMF Staff Papers*, 37(2); June.
- Khan, Mohsin and Knight, Malcolm D. (1981); 'Stabilisation programs in developing countries: a formal framework', *IMF Staff Papers*, 28(1); March.
- Khan, Mohsin and Knight, Malcolm D. (1985); 'Fund-supported adjustment programs and economic growth', *IMF Occasional Paper No. 41*. Washington: IMF; November.
- Khan, Mohsin and Knight, Malcolm D. (1988); 'Import compression and export performance in developing countries', *Review of Economics and Statistics*; May.
- Killick, Tony (ed.) (1984A); *The Quest for Economic Stabilisation: the IMF and the Third World*. London: Overseas Development Institute and Gower Publishing Company.
- Killick, Tony (ed.) (1984B); *The IMF and Stabilisation: Developing Country Experiences*. London: Gower Publishing Company and Overseas Development Institute.
- Killick, Tony (1989); 'Economic development and the adaptive economy', *ODI Working Paper No. 31*. London: Overseas Development Institute; November.
- Killick, Tony and Mweya, F.M. (1990); 'Monetary policy in Kenya, 1967-88', *ODI Working Paper No. 39*. London: Overseas Development Institute; July.
- Loxley, John (1984); *The IMF and the Poorest Countries*. Ottawa: North-South Institute.
- Lynn, Robert and McCarthy, F. Desmond (1989); 'Recent economic performance of developing countries', *Working Paper WPS 228*, Washington: World Bank; July.
- Martin, Matthew (forthcoming); *The Crumbling Facade of African Debt Negotiations*. London and New York: Macmillan and St Martins.
- Moore, Will H. and Scarritt, James R. (1990); 'IMF conditionality and polity characteristics in Black Africa: an exploratory analysis', *Africa Today*, 37(4).

- Mosley, Paul; Harrigan, Jane; and Toye, John (1991); *Aid and Power: The World Bank and Policy-Based Lending* (2 vols.). London: Routledge.
- Pastor, Manuel Jr (1987); 'The effects of IMF programs in the Third World: debate and evidence from Latin America', *World Development*, 15(2); February.
- Reichmann, Thomas M. (1978); 'The Fund's conditional assistance and the problems of adjustment: 1973-75', *Finance and Development*, 15(12); December.
- Reichmann, Thomas M. and Stillson, Richard D. (1978); 'Experience with programs of balance of payments adjustment: stand-by arrangements in the higher credit tranches, 1963-72', *IMF Staff Papers*, 25(2); June.
- Reisen, Helmut and Van Trotsenburg, Axel (1988); *Developing Country Debt: the Budgetary and Transfer Problem*. Paris: OECD.
- Rieffel, A. (1985); 'The role of the Paris Club in managing debt problems', *Essays in International Finance No. 161*, Princeton University Press.
- Sidell, Scott R. (1988); *The IMF and Third World Instability: Is There a Connection?* Basingstoke: Macmillan.
- Taylor, Lance (1988); *Varieties of Stabilization Experience*. Oxford: Clarendon Press.
- World Bank (1989); *Adjustment Lending: An Evaluation of Ten Years of Experience*. Washington: World Bank.
- Zulu, Justin B. and Nsouli, Saleh M. (1985); 'Adjustment Programs in Africa: The Recent Experience', *IMF Occasional Paper No. 34*. Washington: International Monetary Fund.

**Overseas Development Institute
Regent's College
Inner Circle
Regent's Park
London NW1 4NS
England
Telephone: 071-487 7413
Fax: 071-487 7590
Telex: 94082191 ODIUK**