IMF WORKING PAPER

MASTER FILES
ROOM C-130

0440

WP/88/113 INTERNATIONAL MONETARY FUND

Research Department

The Macroeconomic Effects of Fund-Supported Adjustment
Programs: An Empirical Assessment

Prepared by Mohsin S. Khan*

December 23, 1988

Abstract

This paper reviews the existing evidence on the macroeconomic effects of Fund-supported adjustment programs, and provides new estimates of these effects for 67 developing countries with programs during 1973-86. The empirical analysis indicates that in the short run programs have led to an improvement in the current account, the balance of payments, and inflation, but this was accompanied by a decline in the growth rate. In the longer run the positive effects of programs on the external balance and inflation are strengthened, and the adverse growth effects reduced. These results are more definitive than those from previous studies.

JEL Classification Numbers:
1331; 2120

Contents   Page

I. Introduction

The broad objectives of a Fund-supported adjustment program are the attainment of a viable balance of payments, improved long-term growth performance, and low inflation. Generally, the need for an adjustment program, whether supported by the Fund or otherwise, arises when a country experiences an imbalance between aggregate domestic demand and aggregate supply, which is reflected in a worsening of its external payments position and a rise in domestic prices. While supply shocks can cause divergence between aggregate demand and supply, often such imbalances can be traced to inappropriate policies that expand aggregate demand too rapidly relative to the growth of productive capacity of the economy. If this relative expansion of domestic demand is allowed to persist for an extended period the country would experience a widening current account deficit, a loss of international competitiveness and increased distortions in relative prices resulting from higher inflation, a declining growth rate, and a heavier foreign debt burden.

The primary role of the Fund in these circumstances is to assist a country in designing a policy package that includes measures to restore a sustainable balance between aggregate demand and supply, and to simultaneously expand the production of tradables. Such an adjustment program takes the form of a set of policy intentions by the government that is judged by the Fund to warrant financial support. The choice of policies and the policy mix in a program result from extensive negotiations between the country authorities and the Fund, and thus reflect the particular economic situation of the country and the preferences of the government.

A question that is frequently raised in connection with Fund programs is whether such programs have been effective in achieving their macroeconomic objectives. Some writers have argued that, at best, Fund programs do little in the way of improving the economic picture, [1] while others have gone as far as to say that programs worsen the situation by inducing stagflation. [2] Providing a clear-cut answer to this question turns out to be no easy task. There is at present little agreement in the profession either about how to estimate the macroeconomic effects of programs, or about what impact past programs of the Fund have actually had on macroeconomic variables. Despite the fact that there have been a number of studies on the subject over the past decade, [3] one cannot say with certainty whether programs "work" or not. The question is apparently still open.

---

[1] See, for example, Killick (1984).

[2] Taylor (1981) is the leading proponent of this view.

[3] By our count there have been at least 12 cross-country studies of Fund programs published since 1978.

This paper has a two-fold purpose: first, to review the evidence that is available on the effects of programs, paying special attention to the methodologies employed in the various studies that have been produced in the last ten years or so. Second, to estimate the effects of programs on the main macroeconomic variables--the balance of payments, the current account balance, inflation, and growth--in a group of 67 developing countries that implemented programs supported by the Fund during the period 1973-86. This survey of the existing literature, in combination with the new empirical evidence, essentially summarizes the current state of thinking on the topic, and allows one to identify the problems and gaps that remain in evaluating the macroeconomic effects of programs.

The rest of the paper is organized as follows. Section II reviews existing empirical studies on the effects of Fund programs, and analyzes alternative approaches that have been employed in estimating program effects with a view to assessing their relative strengths and weaknesses. The empirical estimates of program effects that are undertaken are discussed in Section III. The concluding section summarizes the main points of the paper and identifies directions for future work on the subject of program evaluation. Appendix I provides a formal description of the alternative methods for estimating the macroeconomic effects of adjustment programs and the econometric biases associated with them, and Appendix II contains the relevant information on the sample of programs and the data utilized in this study.

II. Empirical Studies of the Effects of Programs

It is a popular misconception that the analytical model used to design Fund programs is simply a variant of the well-known monetary approach to the balance of payments. 1/ Consequently, it is argued that the Fund's approach to economic stabilization, generally referred to as "financial programming," relies solely on domestic credit restraint, possibly supplemented by devaluation, to achieve a desired improvement in the balance of payments. If this were indeed the case then one would only have to test the effects of two policies--a reduction in the rate of domestic credit expansion and devaluation--on a single target--the balance of payments--to judge the effectiveness of Fund programs.

However, while it is true that the monetary approach was operationalized in the Fund in the 1950s and 1960s, 2/ and still plays a central role in the design of programs, it does not represent the whole story by a wide margin. Fund programs are complex packages of policy measures

---

1/ See, for example, Dell (1982) and Diaz-Alejandro (1984).
2/ See IMF (1977), (1987).

that have multiple targets--improving the current account balance and the overall balance of payments, raising the growth rate, and reducing infla-tion being the primary ones. Aside from monetary and exchange rate policies, a typical Fund program calls for, inter alia, fiscal measures, such as restraint of government expenditures and increases in taxation, increases in domestic interest rates and producer prices to realistic levels, policies to raise investment and improve its efficiency, trade liberalization, wage policies, and external debt policies. 1/ Naturally, the theory underlying such a policy package combining demand-management, supply-enhancing, and relative-price policies, goes well beyond the relatively straightforward predictions of the monetary approach to the balance of payments.

Since it is theoretically and empirically difficult to link all the policy measures contained in a Fund program to the ultimate targets of balance of payments, inflation, and growth, most studies of the effects of programs have attempted to assess the effects of the overall policy package. In this type of approach the precise nature of the underlying economic relationships and the specific policies adopted are not made explicit, and attention is directed solely at determining whether Fund programs have been "effective" in some sense in achieving the broad macroeconomic objectives for which they were formulated. Cross-country studies of Fund stand-by and extended arrangements have been undertaken periodically within the Fund, and also by writers outside the Fund looking at various aspects of Fund-supported adjustment programs. Although there have also been case studies examining individual-country experiences, at present the literature is dominated by cross-country analyses.

This literature suggests that four distinct approaches have been applied to the evaluation of Fund-supported adjustment programs:

a. The before-after approach, i.e., the difference between macro-economic performance under, or after, a Fund program and performance prior to the program.

b. The control-group approach, i.e., the difference between macro-economic performance in countries with Fund programs and performance in a "control group" of nonprogram countries.

c. The actual-versus-target approach, i.e., the difference between macroeconomic performance under the program and the performance specified in its targets.

d. The comparison-of-simulations approach, i.e., the difference between simulated performance of Fund program-type policies and simulated performance with some other set of policies.

1/ For a discussion of the policy content of Fund programs, see IMF (1987).

In what follows, we will group studies according to the approach that was employed by them.

1.   Before-After Approach

In the literature on the effects of Fund programs the before-after approach has been the most popular. The first study to use this approach was that by Reichmann and Stillson (1978). These authors examined a total of 79 Fund-supported programs implemented during 1963-72 and compared the behavior of the balance of payments, inflation, and growth during the two-year periods before and after the implementation of the program. Using nonparametric statistical tests, 1/ they found that there was a significant improvement in the balance of payments in only about a quarter of all programs. In a majority of cases (over 70 percent) there was no significant change in the balance of payments. Of the 29 programs involving countries with high inflation during the program period, the rate of inflation fell in 6 of 11 programs for which there was a notable deceleration in the rate of domestic credit expansion; in the 9 programs which involved a devaluation, there were 5 in which inflation was higher. Finally, growth performance was examined for 70 programs and it was concluded that on balance, Fund programs did not exert adverse effects on growth rates. In 40 percent of programs the growth rate declined after the inception of the program relative to the previous two year's average rate of growth, but at the same time growth was higher in 47 percent of the programs.

A similar procedure was followed by Connors (1979) who examined a total of 31 programs in 23 countries that were adopted during the 1973-77 period. He compared periods of one year before and after the programs. Also using a nonparametric rank test, Connors (1979) concluded that Fund programs had no discernible effects on the ultimate targets--growth, inflation, 2/ and the current account deficit--or on important intermediate targets, such as the ratio of the fiscal deficit to GDP.

The relationship between fiscal variables and Fund program performance was examined in more detail by Kelly (1982). The methodology was primarily of the before-after variety, and in order to take into account possible lags in adjustment, comparisons were made over both one-year and three-year periods. In a sample of 77 programs (covering 33 countries) during 1971-80, Kelly (1982) observed that when using a one-year comparison the fiscal deficit was reduced in 56 percent of the cases, and that the

---

1/   The authors used a variant of the Mann-Whitney U-test for differences in means.
2/   In the 12 countries in which the ceiling on domestic credit was breached, Connors (1979) found that inflation was significantly higher.

current account and fiscal deficits moved in the same direction in 62 percent of the programs examined. This last result was also supported by regression analysis that showed a positive and statistically significant relationship between changes in the fiscal deficit to GNP ratio and changes in the current account to GNP ratio. Furthermore, in about half the cases there was a decline in the average growth rate over a three-year period, and an increase or no change in the other half.

The study by Killick (1984) also attempted to capture the effects of lags by comparing the behavior of the balance of payments, the current account, growth and inflation a year before the program with the behavior both one and two years after the programs. Killick (1984) employed non-parametric statistical tests to gauge the effects of 38 programs covering 24 countries during the period 1974-79. In contrast with other studies Killick (1984) found that the balance of payments and the current account deteriorated irrespective of the time period over which the comparison was made. However, the difference between the pre-program and post-program values of these variables was not statistically significant. Inflation was reduced, 1/ but the effects on growth were ambiguous. In the first year after the program the rate of growth was higher, but by the second year the positive effect was eroded and the net effect was zero. 2/

Zulu and Nsouli (1985) also constructed before-after measures of program effects in their study of 35 programs implemented in 1980-81 for 22 African countries. They found that growth was lower or the same in the year after the program in about 60 percent of the cases. For the current account and inflation targets the split was even--with as many programs showing an improvement as those showing a worsening or no change.

The most recent study using the before-after approach was that by Pastor (1987) for 18 Latin American countries during 1965-81. Using one-year comparisons and on the basis of alternative statistical tests, Pastor (1987) concluded that Fund programs led to a significant improvement in the balance of payments, but that apparently there was no effect on the current account, inflation, or the rate of growth of nominal GDP.

While easy to employ and seemingly objective, the problem with the before-after approach is that it is based on a strict ceteris paribus assumption and will not yield an estimate of the independent effect of programs on macroeconomic outcomes whenever the nonprogram determinants of these outcomes are changing as between the pre-program period and the program period. The fact of the matter is that these nonprogram determinants, ranging from external factors like industrial-country growth rates,

---

1/   This effect was statistically significant.
2/   In neither case, however, was the effect statistically significant.

terms of trade variations, and movements in international interest rates, to domestic factors such as shifts in weather conditions, do change markedly from year to year in the real world. As illustrated in Appendix I, this means that before-after estimates of program effects will typically be: (a) biased, because this approach incorrectly attributes all of the change in outcomes between the pre-program and program periods to program factors; and (b) unsystematic over time, because estimated program effects for a given year will often be dominated by specific nonprogram influences of that year. Thus, for example, if industrial-country growth jumps upward between year t and year t+1, all t+1 programs will look as if they performed very well, while if industrial-country growth falls sharply in some later year, all programs for that later year will look as if they performed poorly.

These shortcomings of the before-after approach make it a poor estimator of the "counterfactual", defined as the macroeconomic performance that would have taken place in the absence of a program. This is a nontrivial drawback because the counterfactual is perhaps the most appealing yardstick against which to assess program performance and the standard most widely employed in economics to define and measure the impact of government intervention. What would have happened in the absence of a Fund program is by no means the only standard against which to judge the performance of a program, but in many instances, it is the most realistic one. However, the crux of the problem is that the counterfactual is not directly observable and must be estimated. The reason why the before-after approach is flawed as an estimator of the counterfactual is that the situation prevailing before the program is not likely to be a good predictor of what would have happened in the absence of the program, given that nonprogram determinants can and do change from year to year. 1/

2. Control-Group Approach

The control-group approach is designed to overcome the inability of the before-after approach to distinguish between program and nonprogram determinants of macroeconomic outcomes. The basic reasoning behind this approach is as follows. Assume program and nonprogram countries are subject to the same nonprogram determinants, i.e., they face the same external environment. Then, so the argument goes, by comparing before-after changes in outcomes in program countries to those in the control group of nonprogram countries, the effects of nonprogram determinants

---

1/ By making a judgmental correction for the influence of nonprogram factors it is possible to improve upon the estimates of the counterfactual that would emerge from a mechanistic application of the before-after approach. Suffice to say, however, that such judgmental corrections are difficult to make, especially when the range of nonprogram factors is wide.

will cancel out—leaving the difference in group performance to reflect only the effects of Fund programs. Put in terms of the counterfactual, the idea is to use the observed performance of nonprogram countries as an estimate of what the performance of program countries would have been in the absence of a Fund program.

The control-group approach was first used in two studies by Donovan (1981), (1982), that analyzed a sample of programs implemented from 1970 to 1980. The control group was taken to consist of all non-oil developing countries, and the comparisons were carried out over one-year and three-year time horizons. In the first study (Donovan (1981)), which covered a sample of 12 programs (for 12 countries) implemented during 1970-76, it was shown that the improvement in the rate of growth of exports was consistently higher for program countries than for all non-oil developing countries. 1/ The increase in the rate of inflation for program countries was about half that of the control group during the first year, and while it rose when three-year comparisons were undertaken, it was nevertheless well below the average increase in the rate of inflation of all non-oil developing countries. The outcome for growth was not as clear cut. In the one-year comparisons there was a sharp improvement in growth in program countries relative to the control group, but in the three-year comparisons growth in program countries fell by more than it did in nonprogram countries.

In Donovan (1982) the sample of programs was expanded to 78, covering the period 1971-80, and the same analysis as in the first study was undertaken. The balance of payments and the current account positions of program countries was found to improve relative to the control group in both the one-year and three-year comparisons. The increase in inflation in program countries was about half that of non-oil developing countries in the one-year comparisons, and fell to one-third in the three-year comparisons. However, in contrast with the results of the first study, the rate of growth of real GDP fell by more than the average decline experienced by non-oil developing countries in the one-year comparisons, but by less in the three-year comparisons.

Gylfason (1987) also used a version of the control-group approach in his study of 32 programs implemented during 1977-79. The reference group included the developing countries that had experienced balance of payments difficulties during 1975-77, and nonparametric statistical tests were used to determine if the behavior of the macroeconomic variables for program countries over a three-year period was significantly different from that in the control group. 2/ The results of these tests were as follows.

1/ Donovan (1981) did not directly examine the effects on the balance of payments and the current account in this study.

2/ Changes in variables were compared over three calendar years—the year before the program, the program year, and the year following the program.

First, there was an improvement in the balance of payments in program countries relative to the outcomes in the control group. Second, inflation in program countries did not fall, but rather remained approximately the same as the average rate of inflation in the group. Third, the growth rate was not significantly affected by the program.

This pattern of results was basically replicated by Pastor (1987) for Latin American countries. Employing the same statistical methodology as Gylfason (1987), Pastor (1987) also found that the balance-of-payments performance of program countries was significantly better than that of nonprogram countries, and that the differences in inflation and growth performance were not statistically significant.

While the control-group approach copes with some of the problems of the before-after approach, it is by no means ideal. The problem is that program countries can and do differ systematically from nonprogram countries prior to the program period, and this matters for performance evaluation. In short, program countries are not randomly selected. Instead, they are adversely selected in the sense of having relatively poor economic performance prior to the program period. This is not surprising, since after all, a basic requirement for Fund financial support is that the country have a balance-of-payments need. This alone suggests that program countries would be expected to have weaker than average external positions when the program was implemented. In any case, non-random selection of program countries means that simple control-group estimates of programs effects may be biased. 1/ Intuitively, the bias occurs because, under non-random selection, the control-group esti-mator attributes differences in outcomes exclusively to program status, when in fact the difference in starting positions itself is a cause of differences in subsequent performance between the two groups. Furthermore, the direction of the bias can go either way. If past economic difficulties signal less serious current difficulties, even in the absence of a program, then the control-group approach will overstate the beneficial effect of a Fund program. Conversely, if past difficulties signal even more serious present difficulties, then the effect of Fund programs will be understated.

There are ways, however, as demonstrated by Goldstein and Montiel (1986), to modify the control-group approach to reduce some of the biases. In brief, the basic idea is to accept the non-random selection of program countries, to identify the specific differences between program and non-program countries in the pre-program period, and then to control for these differences in initial positions in the comparison of subsequent

---

1/ See Appendix I for a formal demonstration of this point.

economic performance. 1/ To make such a modified control-group approach
work, one needs to identify the relevant reduced-form relationships that
link policy instruments to policy targets, and the policy reaction functions
that show how policy instruments change when the state of economy alters.
The former are needed to determine the effects of alternative policies on
the target variables, while the policy reaction functions are needed to
discern what policies would have been chosen, given pre-program conditions,
in the absence of a program. These empirical relationships are not
easily implemented, and there is still likely to remain a serious margin
of error in estimating the counterfactual. But the main point is that the
error is likely to be significantly smaller than under other approaches.

Goldstein and Montiel (1986) applied the modified control-group esti-
mator to a sample of 68 programs for 58 developing countries implemented
during 1974-81. These authors found that program countries systematically
demonstrated weaker performance, i.e., higher inflation, slower growth,
larger current account and overall balance of payments deficits, than non-
program countries in the pre-program period. Adjusting for these pre-program
differences in performances and taking into account the effects of policy
instruments on targets, Goldstein and Montiel (1986) used regression
analysis to estimate the program effects. Two interesting sets of results
emerge from this study. First, there were no statistically significant
effects of programs on the current account and balance of payments, on
the rate of inflation, or on the growth of real output. Second, the
estimated program effects under the modified control-group estimation
were quite different from those obtained with the standard control-group
estimator. The latter indicated an improvement in the current account, a
slight worsening of the balance of payments, a reduction in inflation,
and a rise in the growth rate associated with programs. When the modified
control-group estimator is used the improvement in the current account
ratio disappears, the deterioration in the balance of payments ratio is
magnified, and the favorable outcomes for inflation and growth reversed. 2/

3. Actual-Versus-Targets Approach

Another strand in the literature on program effects compares actual
outcomes for certain key macroeconomic variables to the targets for such
variables specified by the authorities and the Fund at the inception of
the program. This approach has not been as frequently used as the other

---

1/ By selecting the control-group countries to include only those that
had balance-of-payments problems in the pre-program period, Gylfason (1987)
also attempted to adjust for the bias.

2/ It should be noted, however, that the differences in the outcomes
for the variables between program and nonprogram countries were not
statistically significant.

two approaches. 1/ Reichmann (1978), for example, studied 21 programs for 18 countries that were in effect during 1973-75 and compared the outcomes to targets for the balance of payments, inflation, and growth. He found that the balance of payments targets were met or exceeded in nearly two-thirds of the programs. The targets for inflation were, however, exceeded in over half the programs. There was more success with respect to the rate of growth, with 62 percent of programs meeting the targets.

In a similar vein, Beveridge and Kelly (1980) surveyed the fiscal content of all 105 Fund programs approved during 1969-78. Their focus was on intermediate targets--expansion of domestic credit, government revenues and expenditures, and deficit financing--rather than on the final macroeconomic objectives. Nevertheless, the results for the intermediate targets are informative, since achieving these is generally a necessary condition for meeting the ultimate targets for the balance of payments, inflation, and growth. Beveridge and Kelly (1980) showed that almost all programs contained government revenue and expenditure forecasts, and that both actual expenditures and revenues tended to differ from these forecasts. A shortfall in revenues occurred in about 40 percent of the cases, while expenditures were above projected values in nearly 60 percent of the programs. Consequently, the overall fiscal deficit targets were achieved in only about half the programs, as were the domestic credit ceilings. Finally, governments were more successful in meeting domestic nonbank financing limits than foreign financing targets. The former were satis-fied in almost 70 percent of the programs, but foreign financing of the fiscal deficit exceeded the target in over 60 percent of the cases.

Zulu and Nsouli (1985) also analyzed actual outcomes and targets in their study of African programs approved in 1980-81. They found that the current account targets were met in 38 percent of programs, the inflation targets in about 48 percent of the cases, but growth targets were only achieved in less than 20 percent of the programs.

In the actual-versus-target approach success of a program is measured by the extent to which these targets were achieved. However, this approach sheds little light on how the country's macroeconomic performance was affected by a Fund program. Failing to reach program targets is not necessarily synonymous with a program having no independent effect on outcomes. For example, if targets are too ambitious, or if unexpected nonprogram factors intrude in a negative way, actual outcomes may fall short of targets--even though the program may have produced a much better

---

1/ This is most probably due to the fact that it relies on confidential information on targets that only the country authorities and Fund staff have access to.

outcome than would occur in its absence or under some alternative set of policies. Similarly, under-ambitious targets, or positive shocks, would lead to the meeting or exceeding of targets, even if the program produced weaker effects than would occur in its absence. Because the actual-versus-target approach is not amenable to estimation of the counterfactual, it does not allow one to weigh the costs and benefits of the adopted policy course against the alternatives.

4.   Comparison-of-Simulations Approach

Unlike the other three approaches, the comparison-of-simulations approach does not infer program effects from actual outcomes in program countries. Instead, it relies on simulations of economic models to infer the hypothetical performance of Fund-type policies or policy packages and alternative policy packages. If the aim of the exercise is to evaluate the results of a specific Fund-supported adjustment program, then the use of actual program outcomes is indispensable. However, if the purpose is to evaluate the design and effectiveness of Fund programs in general, then examining the likely effects of alternative policy packages can be quite useful and revealing.

Khan and Knight (1981), for example, constructed a small dynamic econometric model and estimated its parameters on a pooled cross-section time-series sample of 29 developing countries, most of which had programs with the Fund. They then investigated the hypothetical effects of a stabilization program which pursued an external balance target using policies that figure prominently in Fund programs, namely domestic credit restraint and reductions in government expenditures. The simulation experiments showed that such a program produced a sharp price deflation in the first year, followed by a temporary burst of inflation as prices rose back to their equilibrium level. Output, on the other hand, contracted sharply in the first year, then rose temporarily above its full-employment level, approaching that equilibrium level gradually over a period of several years.

In a further study, Khan and Knight (1985) extended their simulation analysis to a comparison of alternative policy packages. Specifically, they compared the results for the balance of payments, inflation, and real output growth of a package of demand-management policies (i.e., a once-for-all reduction in the rates of growth of nominal domestic credit and nominal government expenditures, plus a devaluation) with a combined package of demand-management and structural policies (i.e., the above-mentioned demand-management policies plus a set of structural policies that would gradually raise the rate of growth of capacity output). The demand-management package improved the balance of payments almost immediately, but at the cost of a temporarily higher rate of inflation and a short-run reduction in growth. The simulations of the combined package

showed that structural policies could help to partially offset any short-term adverse effects on growth that might result from demand restraint as well as the inflationary consequences of devaluation. Furthermore, the longer-run effects of Fund-type policies on the balance of payments, inflation, and growth were more favorable than the short-run effects.

Broadly speaking, the comparison-of-simulations approach carries three advantages. First, one can draw on a wider body of adjustment experience, since the data base need not be restricted to countries with Fund programs. Second, since one specifies the policy simulations, one does not have to worry that incomplete implementation of policies, which is often a serious problem in Fund-supported programs, will blur the results. In contrast, the approaches that rely on actual outcomes require untangling of effects of policies from the degree of implementation. Last, and probably most significant, the comparison-of-simulations approach, by its very nature, focuses on the relationship between policy instruments and policy targets. As such, it provides better information on how programs work than do approaches that just look at the bottom line, i.e., at the outcomes for policy targets.

There are, however, practical problems with the comparison-of-simulations approach. To use this approach it is necessary to have at hand an empirical model that incorporates the relations between various policies and certain macroeconomic variables. While there have been some modest attempts at building such models for developing countries, 1/ there is as yet no single model available that covers the whole range of policy measures contained in a typical Fund program. Existing formal models are clearly unable to analyze all the questions relating to Fund programs, and in particular, they do not capture the complex ways in which policy variables are related to the ultimate objectives of programs. Even if one had a suitable model to work with one would still have to face up to the critique of ex-ante econometric policy evaluation, i.e., the so-called Lucas critique. Specifically, the parameters in econometric models may not remain invariant to changes in the policy regime, so that it would be incorrect to treat such parameters as fixed across policy simulations. Put in other words, the actual effects of hypothetical policy packages may turn out to be quite different than the simulated results--and in ways that are difficult to know in advance. An additional concern is that, due to credibility factors, the effect of a given policy may be different when it is implemented within the context of a Fund program than when it is implemented outside it. Agents may believe, for example, that policies agreed to with the Fund are more likely to be carried through and thus would be more inclined to change their behavior. Such confidence, or

---

1/ See, for example, the papers surveyed in Khan and Knight (1985).

credibility, effects are automatically captured by the outcome-based approaches, but not necessarily by a model-based approach.

The studies reviewed above have focused on the effects of Fund-supported programs on macroeconomic targets, such as the external balance, growth, and inflation. Mention should also be made of studies that have considered the probable effects of programs on the distribution of income. Rather than provide quantitative estimates of program effects, Johnson and Salop (1980) and Sisson (1986) examine the likely effects of programs on the degree of income inequality in countries. 1/ These two studies reach similar conclusions, namely that Fund programs have not in general worsened income distribution. Insofar as empirical evidence is concerned, Pastor (1987) is the only study that attempts to ascertain the quantitative effects of Fund programs on income distribution. He finds that for 18 Latin American countries the implementation of a program resulted in a fall in the labor share in income relative to both its pre-program level, as well as to the corresponding share in nonprogram countries. The difficulties of trying to assess empirically the effects of programs on poverty are addressed in a recent paper by Heller et al (1988). This study undertook country-case studies for 9 programs that were in place in seven countries in the early 1980s. The authors concluded that in view of the limited evidence available at this stage it would be unwise to draw firm inferences either way on how Fund programs impact on living standards and poverty.

To summarize, the studies on evaluating the effects of Fund programs on the principal macroeconomic objectives yield the following conclusions:

(i) There is frequently an improvement in the balance of payments and the current account, although a number of studies show no effects of programs;

(ii) Inflation is generally not affected by programs; and

(iii) The effects on the growth rate are uncertain, with the studies showing an improvement or no change being balanced by those indicating a deterioration in the first year of a program.

III. Empirical Analysis of the Effects of Programs

In order to take a fresh look at the empirical effects of Fund-supported programs on the balance of payments, the current account balance, inflation, and growth, we applied three alternative estimators to a large sample of programs over the period 1973-1986. These estimators, discussed

---

1/ See also Addison and Demery (1985).

at length in the previous section, were the before-after estimator, the
control-group estimator, and the modified control-group estimator that
corrects for some of the biases present in the other two estimators. 1/
The use of alternative estimators not only enables us to place greater
confidence in the results, but also provides useful information on the
sensitivity of program effects to different estimators. As far as we are
aware, only one other study (Goldstein and Montiel (1986)) has taken a
similar agnostic approach in the choice of estimator to evaluate the
effects of programs.

The data set used in the empirical analysis is the most up-to-date
and comprehensive of those hitherto employed. Out of a total of 74 devel-
oping countries that had upper credit tranche arrangements with the Fund
during 1973-86, our sample contains 67 of these. 2/ In the estimation,
the term "program country" refers to those country-year observations
(288) in which programs were in effect, and "nonprogram country" to the
remaining (650) country-year observations. The latter form the control,
or reference, group against which we compare program performance.

We turn now to specifying the precise tests of program effectiveness,
and to describing the results obtained from these for the sample of pro-
gram under consideration.

1. Specification of tests

The before-after estimator is simply the mean change in the target
variable over some relevant period, usually one-year. If Δy is the
change in the target variable between the program year and the previous
year, and d is a dummy variable for program status, 3/ a convenient way of
calculating the before-after estimator is to run the following regression
on a pooled time-series cross-section sample of program and non-program
countries: 4/

1/ Because of data limitations we were unable to perform the actual-
versus-targets tests, and the absence of a suitable macroeconomic model
(with the necessary parameter estimates) precluded using the the compari-
son-of-simulations approach.

2/ See Table A1 in Appendix II for a list of the countries and programs.
We excluded 7 small countries with programs because of the lack of data
on certain macroeconomic variables.

3/ That is, d takes on a value of unity in those years when a program
was in effect, and a value of zero in other years.

4/ For a more precise definition of the before-after estimator, see
Appendix I.

(1) $\Delta y = \alpha_1 d$

The outcome variables, that is the y's, include the ratio of the balance of payments to GDP (BOP/Y), the ratio of the current account surplus to GDP (CA/Y), the rate of inflation ($\Delta P/P_{-1}$), and the rate of growth of real GDP ($\Delta yr/yr_{-1}$). By estimating equation (1) for the change in each of these variables, we obtain a direct estimate, as measured by the coefficient α_1, of program effects based on the before-after approach.

The control-group estimator can be calculated quite easily by expanding equation (1) to include a constant term, i.e.:

(2) $\Delta y = \alpha_0 + \alpha_1 d$

The estimated value for the parameter α_1 is equal to the difference in mean changes of the target variables for program and non-program countries. 1/ A statistically significant value for α_1 would thus indicate that the change in the target variable for the program country was different from the corresponding change in that variable in nonprogram countries (the control group).

The modified control-group method involves specifying a reduced-form equation linking changes in each target variable to the program dummy, lagged values of all the target variables, lagged values of policy variables, and any external variables that would have an effect on the target variables. For this test we specified the following equation for each target variable:

(3) $\Delta y = \alpha_0 + \alpha_1(BOP/Y)_{-1} + \alpha_2(CA/Y)_{-1}$

$+ \alpha_3(\Delta P/P_{-1})_{-1} + \alpha_4(\Delta yr/yr_{-1})_{-1}$

$+ \alpha_5(\Delta D/D_{-1})_{-1} + \alpha_6 REX_{-1} + \alpha_7(FD/Y)_{-1}$

$+ \alpha_8(\Delta TOT/TOT_{-1}) + \alpha_9 Trend + \alpha_{10}d$

1/   See equation (6) Appendix I.

where, as before, the Δy includes changes in the ratio of the balance of
payments to GDP, in the ratio of the current account to GDP, in inflation,
and in growth, respectively. The policy variables are the percentage
change in domestic credit $(\Delta D/D_{-1})$, the real effective exchange rate
(REX), 1/ and the ratio of the fiscal balance to GDP (FD/Y). While these
three variables are by no means the only policy measures in Fund programs,
they are nonetheless the more important ones. The only external variable
included is changes in the terms of trade $(\Delta TOT/TOT_{-1})$. Other relevant
international variables, such as the growth of export markets, as well as
domestic variables not included in equation (3), are assumed to be captured
by a linear time trend. The effects of programs will again be determined
by the sign and statistical significance of the parameter α_{10}.

2. Results

Equations (1) - (3) were estimated for each of the four target vari-
ables using pooled time-series, cross-section data (938 observations).
The first set of estimates is based on a one-year comparison, that is,
performance in the program year is compared to performance in the pre-
program year. The results are shown in Tables 1 and 2.

In Table 1 we report the results for the before-after and control-
group tests. Using the before-after criterion for program evaluation we
find that the balance of payments improves during the program year in
relation to the previous nonprogram year. However, the magnitude of the
effect is small and not significantly different from zero at the 5 percent
level. The effect of programs on the current account balance turns out
to be quite strong, and the coefficient measuring this is both positive
and highly significant. Finally, although the coefficients are not
statistically significant, the results indicate that both inflation and
growth are lower in the program year than in the previous nonprogram
year. If one were to take only the signs of the coefficients into account
one would conclude from the before-after tests that programs led to
improvements in the balance of payments and the current account balance,
a reduction in inflation, and a decline in the growth rate. However, on
the basis of standard statistical criteria, it would appear that programs
lead to an improvement only in the current account position.

1/ Strictly speaking, REX is an intermediate objective and not a policy
variable. We assume that in the short run a change in the nominal exchange
rate will be reflected in a similar change in the real exchange rate.

Table 1. Before-After and Control-Group Estimates
of Program Effects: One-Year Comparisons*

Dependent Variable	Before-After Estimate	Control-Group Estimate
$\Delta(BOP/Y)$	0.134 (0.50)	0.730 (2.29)
$\Delta(CA/Y)$	1.102 (3.29)	1.755 (4.35)
$\Delta(\Delta P/P_{-1})$	-38.218 (1.27)	-58.108 (1.61)
$\Delta(\Delta yr/yr_{-1})$	-0.026 (0.07)	0.049 (0.11)

*As measured by the coefficient of the program dummy (α_1); t-values in parenthesis.

Table 2. Modified Control-Group Estimates of Program Effects: One-Year Comparisons*

Dependent Variable	$(BOP/Y)_{-1}$	$(CA/Y)_{-1}$	$(\Delta P/P_{-1})_{-1}$	$(\Delta yr/yr_{-1})_{-1}$	$(\Delta D/D_{-1})_{-1}$	REX_{-1}	$(FD/Y)_{-1}$	$(\Delta TOT/TOT_{-1})$	Trend	Program Dummy
$\Delta(BOP/Y)$	-0.568	0.039	-0.001	-0.012	0.001	-0.002	0.052	0.046	-0.194	0.361
	(15.07)	(1.35)	(0.36)	(0.42)	(0.50)	(1.59)	(1.80)	(6.67)	(5.13)	(1.10)
$\Delta(CA/Y)$	-0.122	-0.507	0.001	-0.115	-0.001	-0.001	0.049	0.082	-0.044	0.846
	(2.70)	(14.59)	(0.04)	(3.51)	(0.09)	(0.45)	(1.43)	(9.93)	(0.97)	(2.15)
$\Delta(\Delta P/P_{-1})$	1.173	-1.445	4.548	2.759	-4.099	0.072	-11.982	-0.159	-3.611	-17.808
	(0.49)	(1.82)	(26.52)	(1.73)	(32.65)	(0.78)	(6.63)	(0.37)	(1.51)	(0.86)
$\Delta(yr/yr_{-1})$	0.128	-0.015	-0.001	-0.888	0.002	-0.001	0.073	0.033	-0.096	-0.888
	(2.83)	(0.44)	(0.03)	(27.06)	(0.08)	(0.19)	(2.19)	(4.02)	(2.12)	(2.25)

*T-values in parentheses below coefficients.

The control-group estimations, also shown in Table 1, display broadly the same pattern as was evident in the before-after tests. The balance of payments improvement in program countries relative to nonprogram countries is quite noticeable, and this time the effect is statistically significant at the 5 percent level. The current account improvement of program countries is also larger once allowance is made for the performance of nonprogram countries, and the coefficient is highly significant. Inflation is reduced by a greater amount in program countries, but although the standard error is reduced, the effect remains insignificant. The effect of programs on the growth rate is the only one that differs in sign from the before-after estimations. In the control-group case we observe that programs are associated with an increase in the growth rate rather than a decrease. However, one cannot take this difference in results too far as the estimated coefficient was not statistically significant in either the before-after tests or in the control-group tests.

Table 2 shows the results for the modified control-group estimator, that is, equation (3). 1/ First of all we note that adjustment for preprogram differences does appear to matter. The coefficients of the own lagged values of each of the outcome variables are all statistically significant in the respective equations. Furthermore, in three out of the four equations the lagged value of at least one of the other outcome variables shows up as significant. This result demonstrates the statistical importance of pre-program conditions, and therefore suggests that the control-group estimator is biased for our sample of programs. The lagged percentage change in domestic credit turns out to have a significant effect only in the inflation equation (although with an incorrect negative sign), and the lagged real effective exchange rate in none. 2/ The lagged fiscal balance variable yields very plausible results. A decrease in the fiscal deficit improves the balance of payments and the current account, lowers inflation, and raises the growth rate. The coefficient measuring the effects of the fiscal balance is also significant in the inflation and growth equations. The change in the terms of trade has positive and significant effects on the balance of payments, the current account balance, and growth, but no statistically significant effect on inflation.

The more relevant results from our standpoint are those dealing with program effects. In Table 2 the coefficient of the program dummy in the balance of payments equation is positive, though unlike in the control-group estimations, it is not statistically significant. The current

---

1/ Country dummies were introduced into the specification to allow for inter-country differences in the dependent variables. The coefficients of these dummy variables are not, however, reported in order to economize on space.

2/ This does not of course imply that the current values of these variables have no effect on the target variables.

account improves in the program year and while this improvement is considerably less than that observed in the before-after and control-group tests, it is nonetheless significant at the 5 percent level. Programs also appear to lead to a reduction in inflation, but this change is not statistically significant. Finally, we find that the growth rate is significantly reduced in program countries relative to the change in nonprogram countries. This last result is broadly consistent with other empirical evidence on the short-run growth effects of Fund programs. 1/

The use of one-year comparisons to evaluate Fund programs can, however, be misleading. Too short a time horizon runs the risk not only of failing to capture the full program effects, but even of misrepresenting the direction of such effects when the short-run impacts are the opposite of long-run effects. Factors such as the downward stickiness of prices and wages, limited mobility of factors of production, and the difference between short-term and long-term price elasticities, can combine to produce a one-year verdict on programs that is very different than, say, a two-year or three-year verdict.

To address this time horizon issue we re-did the estimations comparing the behavior of the target variables in the pre-program year with the average behavior of these variables during the program year and the succeeding year. In other words, the target variable y is defined as the average of its values in periods t and t+1, and the change in y as the difference between this average value and the value in the previous year. This "two-year comparison" is expected to pick up some of the lagged responses of the target variables to the implementation of a program. 2/ For convenience, our focus from now on is solely on the modified control-group estimator, which we have argued is the most appropriate of the three utilized in the empirical tests.

The results of the two-year comparisons are shown in Table 3. Broadly speaking, the results are similar to those obtained in the one-year comparisons reported in Table 2. What is noteworthy is that the program effects are in line with our hypothesis regarding the lengthening of the time horizon. We find that the coefficient measuring program effects on the balance of payments almost doubles in size from that obtained in the one-year comparisons, and also that it becomes statistically significant. There is a substantial increase of the program coefficient in the current account equation as well, and the statistical significance is maintained.

1/ See Khan and Knight (1985).

2/ While, in principle, one could also undertake three-year comparisons as well, the way the tests are set up would require assuming that policies were unchanged for three years (see Appendix I). This is obviously not very plausible.

Table 3. Modified Control-Group Estimates of Program Effects: Two-Year Comparisons*

Dependent Variable	$(BOP/Y)_{-1}$	$(CA/Y)_{-1}$	$(\Delta P/P_{-1})_{-1}$	$(\Delta yr/yr_{-1})_{-1}$	$(\Delta D/D_{-1})_{-1}$	REX_{-1}	$(FD/Y)_{-1}$	$(\Delta TOT/TOT_{-1})$	Trend	Program Dummy
$\Delta(BOP/Y)$	-0.539 (15.98)	-0.005 (0.19)	-0.002 (1.02)	-0.025 (1.01)	0.002 (1.13)	-0.002 (1.57)	0.027 (1.07)	0.034 (5.52)	-0.203 (6.01)	0.699 (2.38)
$\Delta(CA/Y)$	-0.049 (1.21)	-0.660 (21.07)	-0.001 (0.48)	-0.097 (3.29)	0.001 (0.33)	-0.002 (0.11)	0.032 (1.04)	0.051 (6.92)	-0.010 (0.25)	1.211 (3.41)
$\Delta(\Delta P/P_{-1})$	1.036 (0.49)	0.992 (0.61)	1.825 (11.94)	-0.453 (0.29)	-2.127 (19.02)	0.065 (0.78)	-13.798 (8.57)	-0.190 (0.49)	0.317 (0.15)	-21.129 (1.14)
$\Delta(\Delta yr/yr_{-1})$	0.077 (2.26)	0.013 (0.52)	0.003 (1.20)	-0.989 (40.40)	-0.002 (0.99)	-0.002 (1.53)	0.086 (3.37)	0.033 (5.39)	-0.140 (4.13)	-0.644 (2.19)

*T-values in parentheses below coefficients.

While program effects on inflation continue to be statistically insigni-
ficant, one does observe an increase in the (absolute) size of the program
coefficient. This would indicate a tendency of programs to reduce inflation
by more when the evaluation period is extended. A similar pattern to that
of inflation emerges in the results of the growth equation. The empirical
evidence confirms that the negative effects of programs on the growth
rate are reduced in the two-year comparisons relative to the one-year
comparisons. The coefficient in the former declines by about 20 percent.

What the results in Table 3 indicate basically is that the positive
program effects on the balance of payments, the current account balance,
and inflation become stronger, and the negative effects on the growth
rate become weaker, as one stretches out the time period over which the
comparisons are made. It is certainly conceivable that further extension
of the time period would yield even stronger positive effects. But a
word of caution is in order here. Just as too short a period can yield
misleading or erroneous results, so can too long a period, as it increases
the likelihood that factors unconnected with programs will influence the
outcomes. The problem is that the choice of the time period over which
performance is judged is essentially arbitrary. For some variables it
may be sufficient to examine performance over one year, while for others
one may need to look over several years. It would, therefore, be useful
in evaluating programs to review the existing evidence on the time lags
associated with various macroeconomic and structural policies. Based on
this evidence the evaluation procedure could be appropriately designed
to capture the total effects of programs.

The last question we address here is whether program effects are
different between the 1970s and 1980s. As Fund policies and programs are
continually evolving and adapting to changing circumstances, one would
suspect that this would be reflected in the effects that programs have
on macroeconomic variables. To answer the question we divided the sample
into two sub-periods--1973-79 and 1980-86--and estimated the modified
control-group equation (with two-year comparisons). The results for the
coefficient of the program dummy for each sub-period are reported in
Table 4.

It is readily apparent from Table 4 that there are indeed substantial
differences in program effects between the two sub-periods. In the 1973-79
sub-period the program effects on the balance of payments was very small
and statistically insignificant. However, in the 1980s programs had a far
larger positive effect on the balance of payments, and also this effect
was significantly different from zero. The beneficial effect of programs
on the current account, however, is reduced in the 1980-86 period as compared
to the earlier period, but the effects in both cases are statistically

Table 4. Modified Control-Group Estimates of Program
Effects for Sub-Periods*

Dependent Variable	1973-79	1980-86
$\Delta(BOP/Y)$	0.047 (1.09)	0.848 (2.38)
$\Delta(CA/Y)$	1.641 (2.42)	0.748 (2.01)
$\Delta(\Delta P/P_{-1})$	11.280 (2.92)	-21.155 (0.69)
$\Delta(\Delta yr/yr_{-1})$	-0.353 (0.63)	-0.282 (0.80)

*As measured by the coefficient of the program dummy (α_{10});
t-value in parenthese.

significant. The most striking difference to emerge is that for infla-
tion. According to the estimates in Table 4, programs had a significantly
positive effect on inflation in the 1970s. The sign of the program co-
efficient is reversed in the 1980-86 period, although it will be noted
that the coefficient is not statistically significant. Finally, we
come to the growth results. The estimated coefficient of the program
dummy in both sub-periods is not statistically significant--in contrast
with the estimates for the entire period--but the relative sizes of the
coefficient indicate that programs in the 1980s have less adverse effects
on the growth rate. This particular result would be consistent with the
view that recent Fund programs have placed more emphasis on growth, and
accordingly have employed a variety of structural policies towards this
end. As structural policies are not represented in the basic equation
specified here (equation (3)), the growth-orientation of Fund programs in
the 1980s is very likely being captured by the coefficient of the program
dummy.

In summary, on the basis of the tests here, and in particular the
modified control-group tests, we can form some conclusions for Fund
programs implemented during the 1973-1986 period. As a result of the
programs:

(i) There was an improvement in the balance of payments, but this
improvement was statistically significant only when the time period of
evaluation was extended beyond the program year. This result held both
for the entire period of analysis as well as for the 1980s.

(ii) The current account deficit was reduced. This result was the
most robust of all, and stood up irrespective of changes in the time
period over which performance was measured or over which the analysis was
conducted. On average, the implementation of a program led to about a
1 percentage point improvement in the ratio of the current account to GDP.

(iii) The inflation rate was lowered, but this reduction did not show
up as statistically significant in any of the tests performed.

(iv) The growth rate declined in the program year, but when the time
horizon of performance evaluation was extended beyond the program year
the adverse growth effects diminished. Furthermore, there is evidence
the 1980s programs had less negative effects on growth, reflecting perhaps
the greater emphasis being placed on growth-enhancing policies in Fund
programs in more recent years.

V. Conclusions

Over the years there have been a number of empirical studies examining the effects of Fund-supported adjustment programs on key macroeconomic variables, like the current account and the overall balance of payments, inflation, and the rate of economic growth. Such evaluations play an important role in the design of programs as the lessons they yield--positive and negative--can be, and have been, incorporated into the thinking and operations of the Fund. Given the important role of the Fund in the adjustment process of developing countries, examination of the programs supported by the Fund are of obvious interest to the international community as well. The objective of this paper was to survey the studies produced both within the Fund as well as outside the institution, with a view to assessing where we stand at present with respect to our knowledge about the effectiveness of past programs, and further to provide new estimates of the effects of programs using a comprehensive data set covering most of the programs implemented during 1973 to 1986.

The survey of the literature on the macroeconomic effects of Fund-supported programs would seem to point to two broad conclusions.

First, the empirical analysis that is available has been conducted using different methodologies, the relative merits of which deserve careful scrutiny. Most of the studies have attempted to gauge program effectiveness by comparing macroeconomic outcomes in program countries with the targets of these programs, with the performance prior to the implementation of the program, or with the observed performance of countries without programs. If, as we have argued in this paper, the proper standard for measuring program effects is to compare the macroeconomic outcomes under a program with the outcomes that would have emerged in the absence of a program, or under a different set of policies, then none of the approaches generally employed so far--the before-after, the control-group, and the actual-versus-targets approaches--is fully satisfactory. More recently, there have been some limited attempts to apply the so-called counterfactual criteria through estimation of policy reaction functions for program and nonprogram countries and through simulation experiments with macroeconomic models. One can place more confidence in the results yielded by these later studies.

Second, one would be hard-pressed to extract from existing studies strong inferences about the effects of Fund-supported adjustment programs on the principal macroeconomic targets. There is some apparent consistency in the case of the results for the current account and the overall balance of payments--with a small majority of studies indicating that programs lead to an improvement in both these variables--but the picture for inflation

and growth performance is very unclear. On the basis of existing studies, one certainly cannot say whether Fund programs lead to an improvement in in inflation and growth performance. In fact, it is often found that Fund programs are associated with a rise in inflation and a fall in the growth rate. By and large the evidence is fairly inconclusive.

More definitive, albeit still not conclusive, evidence emerges from the independent tests performed in this paper. First of all it is clear that for our sample both the before-after and simple control-group estimates are rejected in favor of the results of the modified control-group approach. Second, and more importantly, we find, using the pre-ferred estimator, that in the first year of the program there is typically an improvement in the balance of payments and the current account balance, and a reduction in the rate of inflation. However, programs do involve some cost in terms of a decline in the growth rate. Basically these results are in line with the predictions of the theoretical models that underlie the design of Fund programs (IMF (1987)). In the short run the demand-reducing elements of the policy package dominate--thereby depressing output while improving the balance of payments and inflation pictures. Supply-enhancing policies start to take effect after some lag so that we do not observe any positive growth effects of programs when making one-year comparisons. Support for this hypothesis is provided by our comparisons of performance in the pre-program period with the average performance in the program year and the succeeding year. We found indeed that the positive effects of programs on the balance of payments, the current account, and inflation were strengthened, and the negative effects on growth were reduced when the evaluation period was lengthened. Programs also appear to have been more effective in improving the external balance in the 1980s as compared to the 1970s, and the costs have been relatively smaller.

Basically these new empirical results indicate that, on average, Fund-supported programs have been more "effective" than previous analyses on the subject would suggest. However, an important drawback of our approach is that no account is taken of the degree of implementation of the policies agreed to between the Fund and the country. In other words, all programs are treated alike, whether or not the policy intentions contained in the program were carried out. By mixing countries with varying implementation records in our sample we may well be biasing the judgment about effective-ness. Had the tests been restricted to only those countries that success-fully implemented the recommended policies, it is conceivable that an even more positive picture would emerge.

The question is then where we go from here in terms of future evalu-ation work. The principal message of this paper is that comparing the macroeconomic outcomes of a program with the corresponding outcomes ob-tained under an alternative set of feasible policies is the most appropriate

way of judging the effects of programs. However, the difficulties involved in following this approach should not be downplayed. Criteria based on the determination of the counterfactual involve a great deal of subjectivity and are difficult to employ in practice. Until we are able to develop suitable techniques to estimate the counterfactual, any study of the quantitative effects of Fund programs has to be necessarily limited in scope. In this context, the modified control-group approach that was utilized in the tests in this paper appears promising and needs to be considered seriously in future evaluation exercises. If one is able to formulate the likely policy reactions of the authorities in the absence of a program, then one will have moved a long way towards estimating the counterfactual.

Generally speaking, more needs to be learned about how economic structure influences the effectiveness of various policies often included in Fund programs. Rather than, say, examining past programs to determine whether devaluation works to improve the trade balance, one might search for the structural characteristics and circumstances that make devaluation more or less effective. These would include, for example, the commodity composition of exports and imports, the nature of the wage-price determination process, and the behavior of supporting expenditure-reducing policies, and so on. With this information at hand one could predict the effects of a devaluation in a particular country, or groups of countries, and then ascertain whether the predictions are satisfied.

The focus on economic structure also raises the difficult issue of the proper level of country aggregation for program evaluation work. In this context there is much to be said for dividing efforts between case studies and larger multi-country studies. Case studies permit one to delve deeply into the specifics of program design and implementation, and to identify special circumstances surrounding the program. [1] On the other hand, it is an extremely time-consuming and expensive way to obtain evidence about program effects, and it may be difficult to generalize from the findings of only a few case studies. Large cross-country samples, in contrast, are easier to prepare and update, are more amenable to the application of uniform statistical techniques, and have the virtue of yielding "representative" information, yet in the process one loses some of the country-specific aspects of programs, including the degree to which policies were implemented. Seen in this light, supplementing the results from cross-country studies with information from case studies that highlight some specific issue may be a convenient compromise.

---

[1] See, for example, the recent case studies of Korea by Aghevli and Marquez-Ruarte (1985), and of Turkey by Kopits (1987) and Saracoglu (1987).

In conclusion, as long as Fund-supported programs are to be an integral part of the adjustment strategies of developing countries, the search for an appropriate way to conduct evaluations of past Fund-supported adjustment programs must obviously continue. This paper provides the basis on which future multi-country evaluation exercises can build upon.

Methodological Issues in Evaluating Fund Programs

This Appendix addresses in a more formal way some of the methodological issues associated with the counterfactual approach to estimating program effects.

Suppose that the j'th target variable in country i is determined according to:

$$(1) \quad y_{ij} = \beta_{oj} + x_i' \beta_j + w' \alpha_j + \beta_{ij}^{IMF} d_i + \epsilon_{ij}$$

where y_{ij} is the target variable (e.g., the balance of payments, the current account, the inflation rate, the growth rate, etc.), x_i is a vector of policy instruments (e.g., domestic credit, the exchange rate, the fiscal deficit, etc.), w is a vector of foreign exogenous variables (e.g., industrial country growth, international interest rates, etc.), d_i is a dummy variable, and ϵ_{ij} is a random shock. The parameter vectors are β_{oj}, β_j and α_j are assumed for simplicity to be uniform across countries. The dummy variable d_i takes on the value 1 if a Fund program is in effect during the period in question and zero otherwise, while the parameter β_{ij}^{IMF} measures the effect of the program during this period on the variable y_{ij}.

It is important to note that this definition of β_{ij}^{IMF} means that x_i refers to the policies that would be adopted in the absence of a program. The vector x_i is therefore directly observable only for nonprogram countries; for program countries the x_i must be estimated. One way in which x_i can be estimated is via the following reaction function:

$$(2) \quad \Delta x_i = \gamma' [y_i^d - (y_i)_{-1}] + \eta_i,$$

where y_i is a vector of target variables, y_i^d is a vector of their desired values, and η_i is a vector of random shocks. The parameter matrix γ' is conformable to Δx_i and y_i, and Δ is a first-difference operator, $\Delta x = x - x_{-1}$. Equation (2) basically says that the change in country i's macroeconomic policy instruments between the current and previous period will be a function of the difference between the desired value of the target variables this period and their actual value last period.

This model can be employed to examine the statistical properties of the before-after and control-group approaches to counterfactual estimation of program effects. The <u>before-after</u> estimator, call it β_{ij}^{BA}, is:

(3) $\beta_{ij}^{BA} = \Delta y_{ij}$ for iεP,

where P denotes the set of program countries during the current period. The expected value of this estimator, conditional on observed values of the foreign exogenous variables, is:

(4) $E\ (\beta_{ij}^{BA}|i\varepsilon P, \Delta w') = \overline{\beta_{ij}^{IMF}} + \overline{\Delta w'}\alpha_j$

$$+ E\ (\Delta x_i'\beta_j + \Delta\ \varepsilon_{ij}|i\varepsilon P,\ \Delta w),$$

which is equal to the true value β_{ij}^{IMF} only if:

(5) $E\ (\Delta x_i'\beta_i + \Delta\varepsilon_{ij}|i\varepsilon P,\ \Delta w) = -\overline{\Delta w'}\alpha_j.$

That is, the before-after estimator is unbiased if one expects that the nonprogram determinants of y_{ij} would have behaved in such a way as to leave y_{ij} unchanged, on average, between the program and nonprogram periods.

The <u>control-group</u> estimator, β_{ij}^{CG}, is given by:

(6) $\beta_{ij}^{CG} = \Delta y_{ij} - \overline{\Delta y_{Rj}},$

where $\overline{\Delta y_{Rj}}$ is the average value of Δy_{ij} over some set R of nonprogram countries. Since we can observe Δx and $\Delta\varepsilon_{ij}$ for all iεR, our information set, call it Ω, now consists of:

$$\Omega = \{(\Delta x_i, \Delta\varepsilon_{ij}\ \text{for}\ i\varepsilon R), \Delta w\ \}.$$

Taking expectations of β_{ij}^{CG}, conditional on iεP and Ω, we have:

$$(7) \quad E(\beta_{ij}^{CG} | i\epsilon P, \Omega) = \beta_{ij}^{IMF} + E(\Delta x'_i \beta_j + \Delta \epsilon_{ij} | ' i\epsilon P, \Omega) -$$

$$(\overline{\Delta x'}_R \beta_j + \overline{\Delta \epsilon}_{Rj}).$$

The control-group estimator will be unbiased if:

$$(8) \quad E(\Delta x'_i \beta_j + \Delta \epsilon_{ij} | i\epsilon P, \Omega) = \overline{\Delta x}_R \beta_j + \overline{\Delta \epsilon}_{Rj}.$$

In other words, if it can be expected that, in the absence of the program, the program country would have behaved just like the average member of the nonprogram reference group. Although the control group estimator (unlike the before-after estimator) controls for the effects of changes in the global economic environment on target variables--because such global factors are assumed to affect program and nonprogram countries equally--it introduces a new source of bias, namely the characteristics of nonprogram countries. If nonprogram countries differ systematically from program countries prior to the program period--in ways that matter for subsequent economic performance--then B_{ij}^{CG} will be biased.

An alternative to the before-after and control-group estimators can be derived by using equation (2) to substitute out the unobservable policy changes that would occur in the absence of a Fund program (i.e., for x_i) from equation (1). The <u>modified control-group</u> estimator is:

$$(9) \quad y_{ij} = \beta_{oij} - (y_{ij})_{-1}'\gamma\beta_j + (x'_i)_{-1}\beta_j + w'\alpha_j +$$

$$\beta_j^{IMF} d_i + (\epsilon_{ij} + \eta_i \beta_j),$$

where the desired values y_{ij}^D have been subsumed into the constant.

Econometric estimation of (9) produces an estimate of β_j^{IMF} which is not subject to the criticisms leveled at the before-after and control-group estimators above. This equation takes care of the estimation of the counterfactual by controlling for the factors that are systematically related to the policies that would have been followed in the country without the program--i.e., the lagged values of target variables and policy instruments.

Data Description

1. Programs

The empirical analysis in Section III was conducted for a sample of 288 upper credit tranche programs (Extended Fund Facility and Stand-By arrangements) implemented for 67 countries during the period 1973-86. The countries are listed in Table Al, with the program years indicated by an asterisk. The total number of observations in the sample (938) thus comprises 288 program-years and 650 nonprogram-years. The 7 countries excluded from the total sample of all programs were: Belize, Dominica, Equitorial Guinea, Grenada, Guinea, Lao P. D. Republic, and Solomon Islands.

2. Data Sources

All data except for the Real Effective Exchange Rate (REX) were taken from the World Economic Outlook (WEO) datafile. The REXs are calculated as trade-weighted indices of relative consumer price indices expressed in a common currency.

The definitions of the variables are:

BOP/Y   =   balance of payments as a percentage of nominal GDP.

CA/Y   =   current account balance as a percentage of nominal GDP.

$\Delta P/P_{t-1}$ =   percentage change in the consumer price index.

$\Delta yr/yr_{t-1}$ =   percentage change in real GDP.

d   =   program dummy; equal to 1 in program years, and zero
                  elsewhere.

$\Delta D/D_{t-1}$ =   percentage change in domestic credit of the banking system.

REX   =   real effective exchange rate.

FD/Y   =   fiscal surplus as a percentage of nominal GDP.

$\Delta TOT/TOT_{-1}$ =   percentage change in the terms of trade (with
                  the terms of trade defined as the ratio of the
                  unit value of exports to the unit value of imports).

Trend   =   linear time trend for each country.

TABLE A1

FUND PROGRAMS : 1973—86

	1973	74	75	76	77	78	79	80	81	82	83	84	85	86
1 AFGHANISTAN	*													
2 ARGENTINA			*	*			*				*	*	*	*
3 BANGLADESH			*				*	*			*		*	*
4 BARBADOS											*	*		
5 BOLIVIA	*						*							*
6 BRAZIL											*	*		
7 BURMA		*			*	*			*					
8 BURUNDI														*
9 CENTRAL AFRICAN REP.									*		*	*	*	*
10 CHILE		*	*								*	*	*	*
11 CONGO							*							*
12 COSTA RICA								*	*	*			*	
13 COTE D'IVOIRE									*	*	*	*	*	*
14 DOMINICAN REPUBLIC											*		*	
15 ECUADOR											*		*	*
16 EGYPT					*	*								
17 EL SALVADOR										*				
18 ETHIOPIA									*					
19 GABON						*		*	*	*				*
20 GAMBIA, THE											*	*		*
21 GHANA							*				*	*	*	*
22 GUATEMALA											*			
23 GUYANA						*	*	*	*					
24 HAITI			*	*		*				*	*			
25 HONDURAS							*		*	*				
26 HUNGARY										*		*		
27 INDIA								*	*	*	*			
28 INDONESIA	*													
29 ISRAEL			*	*										
30 JAMAICA	*				*	*	*		*	*	*	*	*	*
31 KENYA			*				*	*		*	*	*	*	
32 KOREA			*					*	*		*	*	*	*
33 LIBERIA								*	*	*	*	*		
34 MADAGASCAR								*	*	*	*		*	*
35 MALAWI							*	*	*	*	*	*	*	

FUND PROGRAMS : 1973–86
(CONTINUED)

	1973	74	75	76	77	78	79	80	81	82	83	84	85	86
36 MALI										*	*	*	*	*
37 MAURITANIA								*	*				*	*
38 MAURITIUS							*	*	*		*		*	*
39 MEXICO					*					*	*	*	*	*
40 MOROCCO								*	*	*	*	*	*	*
41 NEPAL													*	
42 NICARAGUA							*							
43 NIGER											*	*	*	*
44 PAKISTAN	*				*			*	*	*				
45 PANAMA						*	*	*	*	*	*	*	*	*
46 PERU					*	*	*	*		*	*	*		
47 PHILIPPINES	*			*			*	*	*		*	*	*	*
48 PORTUGAL						*					*	*		
49 ROMANIA					*				*	*	*			
50 SENEGAL								*	*	*	*	*	*	*
51 SIERRA LEONE							*		*			*		*
52 SOMALIA								*	*	*	*		*	
53 SOUTH AFRICA				*						*				
54 SRI LANKA		*			*		*		*		*			
55 SUDAN	*						*	*		*	*	*		
56 TANZANIA			*					*						*
57 THAILAND										*	*		*	*
58 TOGO							*	*	*		*	*	*	*
59 TUNISIA														*
60 TURKEY						*	*	*	*	*	*	*		
61 UGANDA									*	*	*			
62 URUGUAY											*		*	*
63 WESTERN SAMOA							*				*	*	*	
64 YUGOSLAVIA								*	*	*		*	*	
65 ZAIRE					*		*	*	*		*	*	*	*
66 ZAMBIA				*		*	*		*		*	*		*
67 ZIMBABWE											*			
NUMBER OF PROGRAMS	7	3	7	6	10	11	22	26	32	30	41	30	31	32 = 288

References

Addison, Tony, and Lionel Demery, "Macroeconomic Stabilization, Income Distribution and Poverty: A Preliminary Survey," Overseas Development Institute, Working Paper No. 15 (February 1985).

Aghevli, Bijan B., and Jorge Marquez-Ruarte, A Case of Successful Adjustment: Korea's Experience During 1980-84, IMF Occasional Paper No. 39, (Washington: International Monetary Fund, 1985).

Beveridge, W.A., and Margaret R. Kelly, "Fiscal Content of Financial Programs Supported by Stand-By Arrangements in the Upper Credit Tranches, 1969-79," IMF Staff Papers (June 1980), pp. 205-49.

Connors, Thomas A., "The Apparent Effects of Recent IMF Stabilization Programs," International Finance Discusssion Paper 135 (Washington: Board of Governors of the Federal Reserve System, April 1979).

Dell, Sydney, "Stabilization: The Political Economy of Overkill," World Development (August 1982), pp. 597-612.

Diaz-Alejandro, Carlos, "IMF Conditionality : What Kind?" PIDE Tidings (January-February 1984), pp. 7-9.

Donovan, Donal J., "Real Responses Associated with Exchange Rate Action in Selected Upper Credit Tranche Stabilization Programs," IMF Staff Papers (December 1981), pp. 698-727.

_____, "Macroeconomic Performance and Adjustment Under Fund-Supported Programs: The Experience of the Seventies," IMF Staff Papers (June 1982), pp. 171-203.

Goldstein, Morris, and Peter J. Montiel, "Evaluating Fund Stabilization Programs with Multicountry Data: Some Methodological Pitfalls," IMF Staff Papers (June 1986), pp.304-44.

Gylfason, Thorvaldur, "Credit Policy and Economic Activity in Developing Countries with IMF Stabilization Programs," Princeton Studies in International Finance, No.60, August 1987.

Heller, Peter S., A. Lans Bovenberg, Thanos Catsambas, Ke-Young Chu, and Parthasarathi Shome, The Implications of Fund-Supported Adjustment Programs for Poverty: Experiences in Selected Countries, IMF Occasional Paper No.58 (Washington: International Monetary Fund, 1988).

International Monetary Fund, The Monetary Approach to the Balance of Payments (Washington, 1977).

_____, Theoretical Aspects of the Design of Fund-Supported Adjustment Programs, IMF Occasional Paper No. 55 (Washington: International Monetary Fund, 1987).

Johnson, Omotunde, and Joanne Salop, "Distributional Aspects of Stabilization Programs in Developing Countries," IMF Staff Papers (March 1980), pp. 1-23.

Kelly, Margaret R., "Fiscal Adjustment and Fund-Supported Programs," IMF Staff Papers (December 1982), pp. 561-602.

Khan, Mohsin S., and Malcolm D. Knight, "Stabilization Programs in Developing Countries: A Formal Framework," IMF Staff Papers (March 1981), pp. 1-53.

_____, Fund-Supported Adjustment Programs and Economic Growth, IMF Occasional Paper No. 41 (Washington: International Monetary Fund, 1985).

Killick, Tony, The Quest for Economic Stabilization: The IMF and the Third World (New York: St. Martins, 1984).

Kopits, George, Structural Reform, Stabilization and Growth in Turkey, IMF Occasional Paper No. 52 (Washington: International Monetary Fund, 1987).

Pastor, Manuel, "The Effects of IMF Programs in the Third World: Debate and Evidence from Latin America," World Development (1987), pp. 249-262.

Reichmann, Thomas M., "The Fund's Conditional Assistance and the Problems of Adjustment: 1973-75," Finance and Development (December 1978), pp. 38-41.

_____, and Richard T. Stillson, "Experience with Programs of Balance of Payments Adjustment: Stand-By Arrangements in the Higher Credit Trances, 1963-72," IMF Staff Papers (June 1978), pp. 293-309.

Saracoglu, Rusdu, "Economic Stabilization and Structural Adjustment: The Case of Turkey," in V. Corbo, M. Goldstein, and M. Khan, (eds.) Growth-Oriented Adjustment Programs (Washington: IMF and the World Bank, 1987).

Sisson, Charles A., "Fund-Supported Programs and Income Distribution in LDC's," Finance and Development (March 1986).

Taylor, Lance, "IS/LM in the Tropics: Diagrammatics of the New Struc-
     turalist Critique," in W.R. Cline and S. Weintraub, (eds.) Economic
     Stabilization in Developing Countries (Washington: Brookings
     Institution, 1981), pp. 465-502.

Zulu, Justin B., and Saleh M. Nsouli, Adjustment Programs in Africa:
     The Recent Experience, IMF Occasional Paper No. 34 (Washington:
     International Monetary Fund, 1985).