DM/85/74

INTERNATIONAL MONETARY FUND

Research Department

Evaluating Fund Stabilization Programs with Multi-Country Data: Some Methodological Pitfalls

Prepared by Morris Goldstein and Peter Montiel*

December 2, 1985

Summary

In the context of the continuing debate on Fund conditionality, many researchers both inside and outside the Fund have attempted to estimate the effects of Fund programs through the use of multi-country samples. A key challenge in such work is to estimate the "counter-factual," i.e., what would have happened in program countries in the absence of programs.

Two approaches have been dominant in tackling this problem. The "before-after" approach implicitly estimates the counter-factual on the basis of pre-program outcomes in the program countries. The "control-group" approach, on the other hand, relies on observed outcomes in non-program countries to estimate what would have happened in program countries in the absence of programs.

Both approaches are subject to potentially serious statistical problems. Since the problems associated with the control-group approach are less well known (at least in the context of evaluating Fund programs), this paper focuses on the potential sources of bias in this methodology. In particular, it is shown that unless the control group is randomly selected, observed outcomes in this group may be systematically different from what would have happened in the program-country group in the absence of programs. Further, there are good a priori reasons for believing that the assignment of countries to program and non-program groups is not random. In fact, the paper presents some empirical evidence, drawn from the experience of the 1970s, to support this view. Furthermore, it is demonstrated that the criteria for selection to program status have been correlated with macroeconomic outcomes during the program period. In this connection, estimates of program effects that control for observable

* We are indebted to many colleagues in the Fund for comments on an earlier draft. Abdel Ismael provided very competent research assistance.

pre-program differences between program and non-program countries turn out to be quite different from those that do not. Application of the control-group methodology to multi-country samples requires the removal of "sample-selectivity bias," which can arise from a variety of sources. On a broader level, the paper demonstrates why considerable caution is needed in attempting to estimate and interpret the effects of Fund programs using multi-country data.

I. Introduction

A noteworthy by-product of the continuing debate on the benefits and costs of Fund conditionality has been the development of a considerable empirical literature on Fund stabilization programs (i.e., on stand-by and extended arrangements). Further, while the first studies of Fund program experience were carried out almost exclusively by Fund staff (e.g., Reichmann and Stillson [1978], Reichmann [1978]), the last seven years have witnessed at least as much quantitative scrutiny of Fund programs from outside the Fund 1/ as from within it. 2/

A common practice in many of these studies is to compare the behavior of one or more key macroeconomic variables (e.g., the current account, the overall balance of payments, the rate of inflation, the growth rate of real output, etc.) "before" the program with its behavior "during" and/or "after" the program. Also, to account for changes in the international economic environment that could alter macroeconomic outcomes independently of the program, it has become increasingly popular (e.g., see Donovan [1982], Gylfason [1983]) to supplement the before-after calculations for program countries with a similar comparison for a reference or "control group" of non-program countries.

The primary purpose of this paper is to present and to discuss a number of methodological problems or pitfalls that can cause true program effects to differ from estimated program effects when either the "before-after "approach or the "control-group" approach is employed. 3/ Particular attention is paid to the inferences that can properly be drawn about the

---

1/ See Connors [1979], Cline and Weintraub [1981], Williamson [1982, 1983], *Chapman and Killick [1982], Killick [1984], and Loxley [1985].*

2/ See Beveridge and Kelly [1980], Beveridge [1981], Donovan [1981], [1982], Kelly [1982], Odling-Smee [1982], Gylfason [1983], and Zulu and Nsouli [1985].

3/ While not all of the previous studies of Fund program experience were aimed at identifying the independent effects of Fund programs, this paper evaluates the "before-after" approach and the "control-group" approach as estimators of such program effects.

independent effects of a Fund program from a comparison of program and non-program countries. More specifically, we attempt to spell out the conditions under which the observed behavior of macroeconomic outcomes in non-program countries can serve as a good predictor of the unobserved behavior of program countries in the absence of a program, and to identify the biases in estimates of program effectiveness if these conditions are not satisfied.

Because the issue of program effectiveness is a wide-ranging and controversial one, it is worthwhile at the outset to indicate three particular caveats relevant to this study. First, we interpret or define "program effectiveness" as the difference between the actual macroeconomic performance observed under a program and the performance that would have been expected to take place in the absence of a program. As noted by Guitian (1981, pp. 36-37), this is only one of at least three possible measuring rods. 1/ Two others, both of which have been employed in some earlier studies, are the difference between actual macroeconomic performance under the program and actual performance prior to the program, and the difference between actual performance under the program and the performance specified in the targets of the program. Obviously, these three alternative performance indicators can yield different verdicts about program effectiveness. Our preference for the first measure, despite its subjective nature, rests on the argument that it is the only one that can provide an estimate of the program's independent effect in the real world where non-program factors (e.g., oil price shocks, varying rates of economic activity in industrial countries, etc.) are also operating on observed macroeconomic outcomes.

A second caveat is that the interpretation of "program effects" in this paper depends critically on our definition of a "program." Specifically, a "program" is defined to be in effect when a country has a formal arrangement with the Fund, and not when a country adopts a "Fund-type" policy package on its own. Under this definition of a "program" and using our preferred definition of "program effectiveness," a program would be judged to have no effect if the country would have adopted the identical set of policies anyway, even though the policies themselves may have substantial impact on the economy and even though these "Fund-type" policies could be better than some "other" set of policies. 2/ It might be argued that it is the effects of Fund-type policies rather than of Fund involvement that is the most relevant issue. To investigate the effects of Fund-type policies, it is not necessary to differentiate between program and non-program countries. Instead, the relevant

---

1/ For a wide-ranging discussion of alternative interpretations of "program effects," and of their relative strengths and weaknesses, see Goldstein [1986].

2/ In this connection, one would also have to account for the effects of Fund involvement on the availability of additional external resources either from the Fund itself or through the catalytic effect of Fund involvement on other lenders. For more on this point, see page 7.

comparison would be between macroeconomic outcomes under Fund-type policies and those under some other set of policies. Program countries would no doubt be included in the data set but they would not be identified as such. Nevertheless, much of the existing empirical literature on Fund stabilization programs does make direct comparisons between program and non-program countries. It is therefore of some interest to identify what can and what cannot be legitimately inferred from such comparisons.

Third, this paper deals exclusively with the methodology of estimating program effects. Specifically, we do not offer our own estimates of Fund program effects in this paper. We do present (in Section IV) some empirical examples of how estimated program effects can differ depending on the methodology used, but these should not be viewed as reliable estimates of program effects themselves. In fact, it is one of the central tenets of this paper that reliable estimates of Fund program effects from multi-country data must await, inter alia, further testing of the issues and pitfalls outlined here. 1/ In this sense the calculations presented in this paper should not alter anyone's view about whether Fund stabilization programs "work"; these calculations do however have implications for the types of evidence that one may want to collect in the future to determine if and how programs work.

The plan of the rest of the paper is as follows. In Section II we introduce a simple but fairly general model of the relationship between macroeconomic outcomes and the presence or absence of a Fund stabilization program. This model not only permits Fund programs to affect macroeconomic outcomes in program countries through a variety of channels but it also permits prior macroeconomic outcomes to affect the probability that a country has a Fund program itself. In addition, the model admits the possibility of stabilizing macroeconomic policy actions in the absence of a Fund program. We then use this model to analyze the conditions under which "true" program effects would equal "estimated" program effects under two short-hand calculations: (i) before-after comparisons of (mean) macroeconomic outcomes for program countries alone; and (ii) before-after comparisons of (mean) outcomes for program countries relative to those for non-program countries. Anticipating what follows, potentially serious estimation biases are found to exist when the "selection" of program countries is non-random and when the determinants of macroeconomic outcomes are correlated with the determinants of Fund program selection.

1/ There are many reasons why the empirical examples in the paper are not reliable indicators of program effectiveness. To mention just three: (i) the paper deals with only one (sample-selection bias) of many potential sources of bias in cross-country estimates (e.g., we ignore bias associated with interdependence between outcomes in program and non-program countries, as well as bias from aggregation across different types of programs); (ii) the paper considers only short-term (one-year) effects of programs; and (iii) the calculations cover only a few of the wide range of policy instruments actually specified in Fund programs.

In Section III, we outline a procedure (the "modified control-group approach") for removing sample-selectivity bias from control-group estimates of the effects of Fund programs when the selection of program countries is non-random. This modified control-group approach is also capable (in principle) of providing information on how total program effects are apportioned as among induced changes in policy instruments, induced changes in behavioral parameters, and general confidence effects. Practical estimation problems associated with the modified control group approach are discussed.

Section IV investigates the empirical relevance of the key aforementioned methodological pitfalls. For this purpose, we utilize a sample of developing countries and of Fund stabilization programs over the 1974-81 period. Estimates of Fund program effects are then compared across three alternative estimators--a before-after comparison of mean outcomes for program countries alone, a before-after comparison of mean outcomes for program countries relative to that for non-program countries, and a reduced-form regression estimate of program effects that controls for revealed pre-program differences between program and non-program countries. The three alternative estimators are demonstrated to produce substantially different estimates of Fund program effects. The paper's conclusions are summarized in Section V.

II. Comparing Alternative Estimators of Fund Program Effects

A simple model of program effects

In this section, we introduce an explicit analytical framework (in the form of a simple four-equation model) for analyzing the effects of Fund stabilization programs. For the purposes of this paper, it was desirable for such a model to have four broad features. First, the model should be general enough that the two dominant existing statistical approaches to ex post program evaluation--namely, the "before-after approach" and the "control-group approach," could be treated as "special" cases of the more general model. In this way, the assumptions implicit in the existing methodologies can be identified and evaluated. Second, the model should incorporate non-program determinants of macroeconomic outcomes of both an international and country-specific nature. Third, given the fact that Fund stabilization programs operate primarily by altering the design and/or stance of macroeconomic policies, the model should indicate the determinants of indigenous changes in macroeconomic policy so that the macroeconomic outcomes expected in the absence of a program can be explicitly defined. Put in other words, we want a model where the policy instruments and not just macroeconomic outcomes are endogenous. Finally, the model should indicate what objective factors, if any, determine the probability that a country will have a Fund program

during a given time period. The reason for treating Fund program status or program-country "selection" as an endogenous variable is that this is the only way to investigate the consequences of systematic differences between program" countries and "non-program" countries. Obviously, if such differences exist prior to a program period, they need to be taken into account in any subsequent comparison of program and non-program countries to the extent that they affect macroeconomic performance; failure to do so would mean that we could be attributing differences in macroeconomic performance as between the two country groups to the presence or absence of a program when in reality the differences may primarily reflect "other" factors.

In equations (1)-(4) below, we set out a model of Fund program effects that contains these basic features:

$$(1) \quad \Delta y_{ij} = \Delta x_i' \beta_{ij} + \Delta W' \alpha_{ij} + \beta_{ij}^{IMF} \Delta d_i + \Delta \varepsilon_{ij}$$

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

$$(3) \quad z_i = [y_i^d - (y_i)_{-1}]' \delta + \pi_i$$

$$(4) \quad d_i = 1 \text{ if } z_i > z^*$$

$$d_i = 0 \text{ if } z_i < z^*$$

In these equations, y_{ij} is the jth macroeconomic "outcome" or "target" variable in country i; x_i is a K-element vector of macroeconomic policy variables that would be observed in country i in the absence of a Fund program; W is an M-element random vector of "world" non-program variables; z_i is a random variable which serves as the index of country-specific characteristics that determines the probability that country i will have a Fund program during a given period; d_i is a "dummy" variable which takes the value 1 if a country has a Fund program and the value 0 otherwise; y_i^d is the "desired" value of the vector y_i; z^* is the threshold value of z_i that divides program from non-program countries; ε_{ij}, η_i, and π_i are unobservable error terms (with zero means and fixed variances) which are serially and (for simplicity) mutually uncorrelated; β_{ij}, α_{ij}, β_{ij}^{IMF}, γ, δ are constants with the appropriate dimensions; Δ is the first difference operator; the subscript -1 indicates the previous period; and the superscript ' denotes the transpose of a matrix.

The variable y_{ij} in equation (1) should be thought of as one of the major targets of a stabilization program, such as the current account, the overall balance of payments, the inflation rate, the real growth rate, etc. 1/ Equation (1) can then be interpreted as positing that the change in this macroeconomic outcome or target variable will be a function of four factors--(i) changes in macroeconomic policy instruments (e.g., the rate of domestic credit expansion, government tax revenues and/or government expenditure, the exchange rate, etc.) that would have occurred in the absence of a program; (ii) changes in "world" economic conditions (e.g., changes in world oil prices or changes in real economic activity in industrial countries); (iii) the total effect of a Fund program if the country has a program in place during that period; and (iv) a host of unobservable shocks that are specific to country i.

A special word of comment is appropriate for β_{ij}^{IMF}--the coefficient that indicates the effect of a Fund program on macroeconomic outcomes. In our view, this coefficient should incorporate at least three channels or avenues by which Fund programs can impact on y_{ij}. First, Fund programs can alter the value of macroeconomic policy instruments from what they would be in their absence. Note, however, that since Δx_i is defined as the change in policy instruments that would occur in the absence of a program, Δx_i is directly observable only for non-program countries; in the case of program countries, Δx_i must be estimated (through equation (2)). In any case, the important implication is that a program can affect y_{ij} by making the actual change in policy instruments different from Δx_i. The second potential channel of program impact is by altering what might be called "the general state of confidence" about the economy of country i. Here, the successful negotiation of a credible program with the Fund may, for example, have a positive effect on private and official capital inflows into country i that may indeed be quantitatively more significant than the financial resources supplied by the Fund itself in support of the stabilization package. This is, of course, like the first channel of impact, an empirical question and suffice it to note here that the measurement of such "confidence impacts" of programs is an extremely difficult task in practice. The third and final channel of potential program impact is by changing the parameters β_{ij} for any given size change in the policy instruments. Put in other words, programs can work not only by say, making monetary and fiscal policies more restrictive than they would otherwise be, but also by improving or reducing the effectiveness of any given stance of policy. The ways

---

1/ Such an interpretation would be consistent with Guitian's [1981] view that the broad objective of Fund-supported stabilization program is "... the restoration and maintenance of viability to the balance of payments in an environment of price stability and sustainable rates of economic growth."

in which behavioral parameters can shift in response to policy changes have been outlined by Lucas [1976] but here it is enough merely to note that programs can change expectations of agents in the economy about the future course of x_i and y_{ij} and these in turn can affect β_{ij}. 1/ The assumptions in equation (1) that unobservable country-specific shocks have zero expected means and are serially uncorrelated imply that, ceteris paribus, a negative shock to say, country i's balance of payments in period t is not expected to be repeated in the next period. In other words, the model contains a "regression to the mean" characteristic for macroeconomic outcomes that provides for some automatic stabilization. Of course, only the data can decide whether such an assumption is consistent with the recent experience of program and non-program countries.

Moving on to equation (2), the basic notion represented there is that the authorities display a systematic "policy reaction" to perceived disequilibria in their macroeconomic target variables. More specifically, equation (2) says that the change in country i's macroeconomic policy instruments as between the current and previous period will be a function of the difference between the desired value of the macroeconomic target variables this period (y_i^d) and their actual value last period $(y_i)_{-1}$, with γ serving as the coefficient that indicates the responsiveness of the policy instruments to such target disequilibria. For example, in the case of stabilizing policy behavior, equation (2) would suggest that a current account deficit last period that was large relative to the authorities' target deficit would call for a downward adjustment in say, the rate of domestic credit expansion this period. Since Δx_i is defined as the change in policy instruments in country i that would occur in the absence of a program, equation (2) spells out "normal" policy behavior by the authorities and thus provides one approach to estimating the "counter-factual" for program countries. Also, so long as γ carries the correct sign, equation (2) also implies that there may well be stabilizing policy action in non-program countries. Idiosyncratic country-specific policy behavior is intended to be captured by the error term (η_i)in equation (2).

Equations (3) and (4) constitute perhaps the greatest departure in this model from the earlier program evaluation literature by suggesting that the presence or absence of a Fund program should itself be treated endogenously, and in particular, as a function of observable country-specific characteristics. There are strong a priori reasons for

---

1/ For example, an announced new real exchange rate target may be viewed as more likely to be adhered to if it is a component of a Fund program than otherwise. In that case, the response of the private sector to a given change in the real exchange rate may depend on whether that change occurs in the context of a Fund program.

believing that Fund program status is not random. A necessary (but not sufficient) condition for the use of Fund resources is that the country display a "balance of payments need." 1/ This implies that among the population of potential claimants for Fund resources, the sample of countries with Fund programs in place at any given time is likely to have displayed less favorable external balance performance prior to the program period itself than the population at large.

As written, equation (3) uses the difference between the "desired" values of macroeconomic target variables this period and their actual values last period to explain the probability that country i will have a Fund program this period. Under the assumption that the desired target values (y_i^d) are constant over time but not necessarily across countries, this reduces to a formulation where the actual values of macroeconomic outcomes in the pre-program period $[(y_i)_{-1}]$ influence program-country selection. Equation (3) should therefore be capable of capturing systematic selection of program countries by the Fund based on "balance of payments need" because the vector $[y_i]_{-1}$ will include pre-program values of country i's external accounts.

Note further that such a specification of equation (3) also deliberately makes for a potentially serious problem. This is because, as written, the pre-program outcomes $(y_i)_{-1}$ help explain "program selection" in equation (3) and also "policy reaction" (Δx_i) in the absence of a program in equation (2). Looking ahead, since both z_i and Δx_i can influence the change in macroeconomic outcomes as between the program and pre-program year, it can be seen that our model sets up the troublesome possibility that it will be difficult to separate program from non-program determinants of Δy_i. As we shall demonstrate later, this problem will be present so long as the determinants of Fund program status (d_i) are correlated with the determinants of Δy_i, be that via Δx_i or via any other variable explaining Δy_i. Also looking ahead, it can be seen that this problem disappears if program selection is "random" since then δ in equation (3) will be zero; that is, it will not be possible to relate program-country selection (z_i), and hence, ultimately d_i, to any observable objective factors.

With the general outlines of this model of IMF program effects in mind, we can next proceed to analyze how "estimated" program effects will differ from "true" program effects under a variety of shorthand estimation techniques.

1/ While there is no explicit formula for judging "balance of payments needs," the three indicators given foremost attention are the actual balance of payments, the level of international reserves, and recent changes in the level of reserves.

The before-after approach

This approach to ex-post program evaluation has been employed both by Fund staff (e.g., Reichmann and Stillson [1978]) and by outside observers (e.g., Connors [1979], Killick and Chapman [1982]). Although these studies utilize multi-country samples, this approach is not necessarily a cross-section technique because the (implicit) parameters estimated are allowed to differ across countries.

Recalling that β_{ij}^{IMF} is the "true" effect of an IMF program on the jth target variable in country i, the before-after approach estimates β_{ij}^{IMF}, call it $\hat{\beta}_{ij}^{IMF,A}$, as:

$$(5) \quad \hat{\beta}_{ij}^{IMF,A} = \Delta y_{ij}, \qquad i \in P$$

Thus, any change in a target variable in a program country (or group of program countries) is attributed exclusively to program effects. The estimate, $\hat{\beta}_{ij}^{IMF,A}$ is sometimes subjected to (non-parametric) statistical tests of significance, and sometimes not.

The fatal flaw of the before-after approach is that it relies on ceteris paribus assumptions that are highly implausible. To see this, let us introduce only the first equation from the general model of program effects:

$$(1) \quad \Delta y_{ij} = \Delta x_i' \beta_{ij} + \Delta W' \alpha_{ij} + \beta_{ij}^{IMF} \Delta d_i + \Delta \varepsilon_{ij}$$

Now suppose that the previous period was one during which there was no Fund program in effect (so $d_i = 0$ for t-1). Then,

$$(6) \quad \Delta y_{ij} = \beta_{ij}^{IMF} + \Delta x_i' \beta_{ij} + \Delta W' \alpha_{ij} + \Delta \varepsilon_{ij} \text{ for } i \in P$$

Using equation (5), the before-after approach then gives:

$$(7) \quad \hat{\beta}^{IMF,A}_{ij} = \beta^{IMF}_{ij} + \Delta x_i' \beta_{ij} + \Delta W' \alpha_{ij} + \Delta \varepsilon_{ij} \quad \text{for } i \in P$$

Taking expectations of (7), conditional on the presence of a Fund program in country i and on observed changes in the world economic environment, we have:

$$(8) \quad E(\hat{\beta}^{IMF,A}_{ij} / i \in P, \Delta W') = \beta^{IMF}_{ij} + E(\Delta x_i' / i \in P, \Delta W') \beta_{ij} + \Delta W' \alpha_{ij} + E(\Delta \varepsilon_{ij} / i \in P, \Delta W')$$

Thus, we see that the before-after approach would produce an unbiased estimate of program effects [i.e., $E(\hat{\beta}^{IMF,A}_{ij} / i \in P) = \beta^{IMF}_{ij}$] if and only if:

$$(9) \quad E(\Delta x_i' / i \in P, \Delta W') \beta_{ij} + \Delta W' \alpha_{ij} + E(\Delta \varepsilon_{ij} / i \in P, \Delta W') = 0$$

In words, an unbiased estimate of program effects would require that the non-program determinants of y_{ij} would have behaved in such a way as to leave y_{ij} unchanged on average as between the pre- and current program period. Reference to the 1973-81 period when large changes in world oil prices, large year-to-year changes in industrial-country real GNP, and significant shifts in real interest rates created serious difficulties for developing countries' external positions (e.g., see IMF [1983], Goldstein and Khan [1982], Khan and Knight [1983]) gives sufficient reason to doubt that $\Delta W' \alpha_{ij}$ would be zero for most program countries, even over short periods. By the same token, the ex post record of money supply growth and fiscal deficits by developing countries during the same period (e.g., see IMF [1983], Gylfason [1983]) generates skepticism that changes in domestic policy instruments would have been such as to just offset, on the average, the effects of external and internal shocks for most program countries. In short, we should expect the ceteris paribus assumption of the before-after approach to be violated in practice. As such, estimates of "program effects" under this approach are likely to be contaminated by non-program factors.

The traditional control-group approach

This second approach, hereafter called "the traditional control-group approach," has a long history in empirical labor economics but appears to have been first applied to analyzing the experience with Fund programs by

Donovan [1982]. 1/ More recently, Gylfason [1983] has adopted a more sophisticated version of it.

This technique, in effect, uses the behavior of a "control group" (a group of non-program countries) to estimate what would have happened in the program group in the absence of programs. Thus, it implicitly assumes that only the program itself distinguishes the group of program countries from the "control group". It is therefore natural to interpret this as a cross-sectional approach. Specifically, in terms of our model, we can drop all the country i subscripts from the coefficients since these are now assumed to be identical across countries. In addition, because we are now dealing with country groups, β_j^{IMF} represents the mean effect of Fund programs on the jth macroeconomic target variable. The equation for Δy_{ij} is therefore now written:

$$(10) \quad \Delta y_{ij} = \Delta x_i' \beta_j + \Delta W' \alpha_j + \beta_j^{IMF} \Delta d_i + \Delta \varepsilon_{ij}$$

Under the control-group approach, β_j^{IMF} is estimated by:

$$(11) \quad \hat{\beta}_j^{IMF,B} = (\overline{\Delta y_j})_P - (\overline{\Delta y_j})_N.$$

where a bar over a variable represents its mean.

To investigate the properties of this estimator, we again take expectations. Applying this procedure to (11) yields: 2/

$$(12) \quad E(\hat{\beta}_j^{IMF,B}) = \beta_j^{IMF} + E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/i \in P] - E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/i \in N]$$

1/ More broadly, many of the methodological issues discussed in this paper have been analyzed earlier in the labor economics literature on "treatment effects;" for example, see Ashenfelter [1978] and Ashenfelter and Card [1984].

2/ To simplify the notation, expectations of group averages in (12) and elsewhere in the paper are expressed in terms of the "representative" member of each group--i.e., we implicitly assume that all members of a group are identical.

From equation (12), it can be seen that the condition for $\hat{\beta}_j^{IMF,B}$ to represent an unbiased estimate of the true program effects (β_j^{IMF}) is that:

$$(13) \quad E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/i \in P] - E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/i \in N] = 0$$

In words, the groups of program and non-program countries have to be drawn from the same population in the sense that the expected value of the change in non-program determinants of y_{ij} must be same for members of both groups. Comparing equation (13) with equation (9), we also note that the control-group approach is not necessarily less restrictive than the before-after approach. Although the control-group approach controls for the effect of changes in the global economic environment [i.e., the term $\Delta W' \alpha_{ij}$ which appears in (9) drops out as a source of bias in (13) because such global factors are assumed to affect program and non-program countries equally], it does so at the expense of introducing a new source of bias--namely, the characteristics of non-program countries, [i.e., the term, $-E(\Delta x_i' \beta_j + \Delta \varepsilon ij)/i \in N$, appears in (13) but not in (9)]. 1/

The foregoing suggests that the choice between the before-after approach and the control-group approach to estimating program effects ought to depend on one's a priori beliefs about similarities between program and non-program countries and about the relationship between domestic and global determinants of Δy_{ij}. Specifically, if program and non-program countries are believed to be quite similar on average, and if the domestic determinants of Δy_{ij} are not believed to offset international influences on Δy_{ij}, then (13) is more likely to be satisfied than (9) and hence, the control-group approach will provide a better (less biased) estimate of program effects than the before-after approach. We next proceed to investigate: (i) the nature of the bias that results under this methodology when the determinants of program-country selection are correlated with the determinants of macroeconomic performance; and (ii) the nature of the biases in the specific, but also probably the most practically relevant case, when both program-country selection and macroeconomic performance depend on macroeconomic performance prior to the program period.

Non-random selection of program countries

To examine these issues it is helpful to introduce the index of unobservable country-specific characteristics (z_i) that regulates the

---

1/ We are indebted to Rusdu Saracoglu for drawing this to our attention.

probability that country i will have a program during any given period. Specifically, we now employ equation (4) from our general model of program effects:

$$(4) \quad d_i = 1 \quad \text{if } z_i > z^* \qquad d_i = 0 \quad \text{if } z_i \leq z^*$$

where z^* is an arbitrary threshold value for z_i. Instead of also introducing equation (3), assume for the moment that $E(z_i) = 0$ and $E(z_i^2) = \sigma_z^2$. Equation (4) says that a country will have a program if its index of country-specific characteristics is greater than z^*; if not, it will not have a Fund program, at least in that period. The probability that a country will have a program is therefore equal to the probability that $z_i > z^*$.

We can now use equation (4) to rewrite the necessary condition, as previously expressed in equation (13), for an unbiased estimator of program effects under the control approach; that is,

$$(13)' \quad E[(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})/z_i > z^*] - E[(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})/z_i < z^*] = 0$$

Recall that both $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ and z_i are random variables. Suppose that the correlation between these two variables is given by ρ_{xz} and that the expected value of $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ is $\Delta x' \beta_j$. We show in the Appendix that if $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ and z_i have a joint normal distribution, then:

$$(14a) \quad E[(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})/z_i > z^*] \begin{cases} > \Delta x' \beta_j & \text{if } \rho_{xz} > 0 \\ = \Delta x' \beta_j & \text{if } \rho_{xz} = 0 \\ < \Delta x' \beta_j & \text{if } \rho_{xz} < 0 \end{cases}$$

$$(14b) \quad E[(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})/z_i \leq z^*] \begin{cases} < \Delta x' \beta_j & \text{if } \rho_{xz} > 0 \\ = \Delta x' \beta_j & \text{if } \rho_{xz} = 0 \\ > \Delta x' \beta_j & \text{if } \rho_{xz} < 0 \end{cases}$$

In words, if $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ and z_i are correlated ($\rho_{xz} \neq 0$), then our expectation of $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ will depend on the value

taken by z_i. This result is intuitive. Suppose, for example, that ρ_{xz} is positive. Then, "relatively large" values of $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ are associated with "relatively large" values of z_i. Thus, if we know that z_i is "relatively large" for some country i but we do not observe $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$, our expectation is that $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ will also be "relatively large" for this country. Likewise, if ρ_{xz} is <u>negative</u>, then "relatively large" values of z_i will be associated with "relatively small" values of $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$, and observing a "large" z_i will lead us to expect a "small" $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$. On the other hand, if $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ and z_i are known to be uncorrelated, then observing a "large" z_i gives us no basis on which to expect $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ to be either particularly large or particularly small for the ith country.

Since program countries are those for which $z_i > z^*$, program countries as a group will exhibit a "relatively large" z_i. Likewise, the representative z for the nonprogram group will be "relatively small". It follows that if $\rho_{xz}>0$ the difference between the program and nonprogram groups concerning the expected change in target variable y_j consists of <u>both</u> the effect of the program on the change in y_j <u>and</u> the difference between the "relatively large" $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ expected for the program group and the "relatively small" $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ expected for the nonprogram group. Since this second component of the expected difference must be positive, the expected difference will <u>exceed</u> the "true" effect of the program on the change in target y_j. A similar analysis establishes that the expected difference will be <u>less</u> than the true program effect when $\rho_{xz}<0$. When $\rho_{xz} = 0$, it remains the case that z_i is "relatively large" for program countries and "relatively small" for nonprogram ones, but this gives us no reason to expect that $(\Delta x_i' \beta_j + \Delta\varepsilon_{ij})$ will be systematically different between the two groups; therefore, the only difference we are justified in expecting is that attributable to the effects of the program. These considerations imply:

$$
\begin{aligned}
(15)\quad E(\hat{\beta}_j^{IMF,B}) &> \beta_j^{IMF} \text{ if } \rho_{xz} > 0 \\
&= \beta_j^{IMF} \text{ if } \rho_{xz} = 0 \\
&< \beta_j^{IMF} \text{ if } \rho_{xz} < 0
\end{aligned}
$$

The relationships in (15) can be derived formally by substituting (14a) and (14b) in (12).

Thus, ρ_{xz} is the crucial parameter in determining the direction of the bias in the control group methodology. Specifically, if the determinants of program selection (z_i) are <u>positively</u> correlated with the determinants of macroeconomic performance that would have occurred in the absence of a

program ($\Delta x_i' \beta_j + \Delta\varepsilon_{ij}$), then the control-group estimate of program effects ($\hat{\beta}_j^{IMF,B}$) will overstate "true" program effects (β_j^{IMF}). Conversely, only if the determinants of program selection are uncorrelated with the determinants of macroeconomic performance ($\rho_{xz}=0$) is the control-group estimator an unbiased indicator of true program effects.

The significance of the preceding analysis is that it permits us to move beyond the vague statement that "if the program and nonprogram groups are different, then the control-group approach will be biased"--a statement which is not correct-to the precise identification of ρ_{xz} as the critical parameter determining both the presence and direction of bias. 1/ In assessing the adequacy of the control-group methodology, the relevant question then is whether there are any reasons inherent in the nature of the problem that would lead us to believe that this correlation (ρ_{xz}) will be nonzero.

Our model embodies precisely such a nonzero correlation. This is because both the determinants of program status in equation (4) and of "normal" policy changes in equation (2) are linear functions of macroeconomic outcomes prior to the program period. To show this formally, first rewrite the model by taking the transpose of equation (2), substituting for $\Delta x_i'$ in (1), and making some small changes of notation. 2/ This gives:

$$(16) \quad \Delta y_{ij} = \beta_{0ij} - (y_i)_{-1}' \gamma' \beta_j + \beta_j^{IMF} \Delta d_i + \tilde{\Delta\varepsilon}_{ij}$$

$$(17) \quad z_i = \delta_{0i} - (y_i)_{-1}' \delta + \pi_i$$

$$(4) \quad d_i = 1 \text{ if } z_i > z^*$$

$$= 0 \text{ if } z_i \leq z^*$$

where:

$$\beta_{0ij} = y_i^{d\prime} \gamma' \beta_j$$

$$\tilde{\Delta\varepsilon}_{ij} = \Delta\varepsilon_{ij} + \eta_i' \beta_j$$

$$\delta_{0i} = y_i^{d\prime} \delta$$

1/ Note that when equation (4) holds but $\rho_{xz} = 0$, the program and nonprogram groups can be quite different without implying the existence of bias in the control group methodology.

2/ We also drop the global variable $\Delta W'$ and its coefficient α_{ij} from equation (1). This was done to simplify the exposition.

In order to determine whether the control-group estimator of program effects will be biased, we again need to examine the correlations betweeen $-(y_i)'_{-1}\ \gamma'\ \beta_j$ and z_i and between $\Delta\tilde{\varepsilon}_{ij}$ and z_i. These will be determined by the signs of:

$$(18a)\quad \mathrm{Cov}[-(y_i)'_{-1}\ \gamma'\beta_j,\ z_i] = \beta'_j\ \gamma\Sigma y_{-1}\ \delta$$

$$(18b)\quad \mathrm{Cov}(\Delta\tilde{\varepsilon}_{ij},\ z_i) = \sigma^2_{\varepsilon}\ \delta_j$$

where Σ_{y-1} is the covariance matrix of $(y_i)_{-1}$, σ^2_{ε} is the variance of ε_{ij}, and δ_j is the jth component of δ. These results assume that ε_{ij}, η_i, and π_i are mutually uncorrelated.

The crucial thing to notice about the covariances portrayed in equations (18a) and (18b) is that they can in general be expected to be non-zero--a finding that implies that the control-group estimator of program effects will be biased. The more interesting issue however is why this estimator turns out to be biased. Our analysis suggests that the determinants of program status will be correlated with the non-program determinants of Δy_{ij} for two reasons.

First, pre-program values of key macroeconomic target variables, $(y_i)_{-1}$, are likely to trigger policy responses (Δx_i) even in the absence of programs, as originally suggested in our "policy reaction function"; in terms of equation (18a), this shows up as a non-zero value of the covariance $\beta'_j\ \gamma\ \Sigma y_{-1}\delta$.

Second, negative transitory shocks in the pre-program period are by their very nature unlikely to recur during the program period (recall that ε_{ij} has an expected value of zero and is assumed serially uncorrelated), with the result that changes in macroeconomic target variables as between the pre- and program period (Δy_i) will display "regression to the mean" with respect to past shocks; in terms of equation (18b), this shows up as a non-zero value for $\sigma^2_{\varepsilon}\ \delta_j$.

We next want to inquire about the direction of this bias. For that source of bias that arises from "regression to the mean," we can provide an unambiguous answer under reasonable assumptions. This is not so for the bias arising from the existence of policy reaction functions; in this latter case, however, we can spell out the conditions necessary for that source of bias to disappear.

Consider the bias arising from "regression to the mean." Since Fund programs are designed to move target variables in the desired direction, we expect the product of β_j^{IMF} and δ_j to be greater than zero (i.e., $\beta_j^{IMF}\ \delta_j>0$). The logic here is that if a below target value of y_{ij} (e.g., the current account surplus) causes a country to "come to the Fund" ($\delta_j>0$), the program will seek to increase actual y_{ij} ($\beta_j^{IMF}>0$); hence, $\beta_j^{IMF}\ \delta_j>0$. Likewise, for those target variables (e.g., the rate of inflation) for which the likelihood of program participation is increased when $(y_i)_{-1} > y_i^d$, then $\delta_j<0$ and we expect $\beta_j^{IMF}<0$; here too, the product, $\beta_j^{IMF}\ \delta_j$ will still be greater than zero. The relevance of $\beta_j^{IMF}\ \delta_j>0$ is that because we know from equation (18b) that the correlation between $\tilde{\Delta}\varepsilon_{ij}$ and z_i carries the same sign as δ_j, we can conclude that "regression to the mean" contributes to a correlation between the determinants of program status (z_i) and the non-program determinants of Δy_{ij}, i.e., ρ_{xz}, which has the the same sign as β_j^{IMF}. From our earlier analysis, especially equation (15), we then know that in these circumstances, the control-group approach will overstate the true effect of a Fund program. In short, if program countries are more likely to have experienced negative temporary shocks in the pre-program period, a comparison of changes in mean macroeconomic outcomes as between program and non-program countries will, under plausible assumptions, overstate the beneficial effect of a program. This is because a negative shock in the pre-program period simultaneously increases the probability of program participation and increases the probability of a positive change in y_{ij} in the program period. Thus, attributing all of this improvement in y_{ij} to a Fund program overstates the program's true independent effect.

The direction of bias arising from the existence of policy reaction functions themselves depends on the characteristics of such functions, which is of course an empirical question. Nevertheless, we can show that the bias will disappear under two conditions.

The first such condition is that $\delta = 0$, that is, when Fund program status can no longer be related to observable country characteristics and when therefore all countries have an equal probability of becoming program countries. In this case, the covariance represented by equation (18a) is zero as long as π_i is uncorrelated with ε_{ij} and η_i.

In this "random selection case," both sources of bias disappear because the original premise of the "control group" approach, namely that program and non-program countries are similar, is satisfied.

The second condition for the policy-reaction bias to disappear is $\gamma = 0$. Again, this would make the covariance represented in equation (18a) equal to zero under our assumption. In words, $\gamma = 0$ means that the policy reactions of the authorities cannot be systematically related to observable characteristics; that is, we would not observe the systematic policy reaction functions represented by equation (2). Note, however, that even when $\gamma = 0$, the bias in the control-group estimator attributable to "regression to the mean" would still remain. This is so because $\delta = 0$ eliminates the improvement in y_{ij} attributable to non-program policy actions but not that improvement attributable to automatic stabilization from reversible country-specific shocks.

To sum up, we have argued in this section that there are strong ex ante reasons for believing that the past procedures used to estimate the effect of Fund programs in multi-country samples are subject to significant sources of statistical bias. Since the non-program determinants of macroeconomic outcomes cannot in general be expected to behave in such a way as to leave these outcomes unchanged from year to year, the potential problems with the "before-after approach" can be readily acknowledged. The problems with the "control group approach" are also important but they are perhaps more subtle. As shown above, comparing mean macroeconomic outcomes as between groups of program and non-program countries will lead to biased estimates of program effects whenever the determinants of program-selection are correlated with the determinants of macroeconomic outcomes that would have occurred in the absence of a program.

III. Obtaining Unbiased Control-Group Estimates of Program Effects Under Non-Random Selection of Program Countries

In this section, we describe a "modified" control-group estimator and show why it is capable of producing unbiased estimates of program effects even when program and non-program countries are "different." Second, we discuss some of the operational problems that would have to be faced in actually implementing this estimator. Third and finally, we show how this estimator could be used to obtain information not just on total program effects but also on how these effects are achieved. It must be emphasized that we describe only the modifications required to control for observable differences between program and non-program countries. Sample-selectivity bias would remain due to unobservable differences between program and non-program countries. Although statistical procedures are available to handle this source of bias, we do not describe them here. Furthermore, the modifications we discuss also cannot handle

other potential biases (e.g., aggregation effects, interdependence between program and non-program countries) that may distort the true effects of programs.

A modified control-group estimator

Consider the following "modified" estimator ($\hat{\beta}_j^{IMF,M}$) for Fund program effects:

$$(19) \quad \hat{\beta}_j^{IMF,M} = (\overline{y_j})_P - (\overline{y_j})_N - (\overline{x_p} - \overline{x_N})' \beta_j$$

Reference back to equation (11) reveals that this modified estimator differs from the traditional control-group estimator in two respects: the modified estimator contains the additional term $-(\overline{x_p}-\overline{x_N})'\beta_j$, and it is specified in level rather than in first-diference form. 1/

To investigate the properties of this estimator, write our basic equation for the jth macroeconomic target variable in country i in level form:

$$(20) \quad y_{ij} = x_i'\beta_j + W'\alpha_j + \beta_j^{IMF} d_i + \varepsilon_{ij}$$

Taking expectations of equation (19), after substituting from (20), we then obtain:

$$(21) \quad E(\hat{\beta}_j^{IMF,M}) = \beta_j^{IMF} + E[(\overline{\varepsilon_j})_P - (\overline{\varepsilon_j})_N]$$
$$= \beta_j^{IMF} + E(\varepsilon_{ij}/z_i>0) - E(\varepsilon_{ij}/z_i \leq 0)$$
$$= \beta_j^{IMF}$$

Thus, we see that the modified control-group estimator will be unbiased so long as the unobservable country-specific determinants of y_{ij}, namely

---

1/ This is only one of several equivalent "modified" estimators that could be proposed. Their common feature is that outcomes are measured net of observable non-program influences that can be estimated on the basis of pre-program information.

(α_{ij}), are uncorrelated with the determinants of program status (z_i). In such a case, one can set $E(\varepsilon_{ij}/z_i>0) = E(\varepsilon_{ij}/z_i < 0) = 0$, and thereby justify the last equality above.

The reason why the modified estimator is unbiased can be explained intuitively using the conclusions from Section II. Recall that we established there that the "traditional" control-group estimator would be biased if the non-program determinants of Δy_{ij}--namely, changes in domestic macroeconomic policy and changes in unobservable shocks--were different systematically as between program and non-program countries, i.e., if Δx_i and $\Delta\varepsilon_{ij}$ were correlated with Fund program status (z_i). The "modified" control-group estimator removes both sources of bias present in the "traditional" version. By substracting the term $(\bar{x}_p-\bar{x}_N)'\beta_j$, an adjustment is made for any differences in indigenous macroeconomic policy between program and non-program countries. Regarding the second potential source of bias (regression to the mean), note that systematic differences between program and non-program countries concerning changes in unobservable shocks are to be expected only because the program-selection rule makes it more likely that countries with unfavorable shocks in the pre-program period will subsequently adopt programs. But also note that the expected level of such shocks, that is, $E(\varepsilon_{ij})$, is zero for all countries. Thus, under our assumptions about the distribution of ε_{ij}, this source of bias is present in estimators expressed in first-difference form which fail to control for prior shocks but not in those (like the modified estimator) expressed in level form.

Operational aspects of the modified control-group estimator

The traditional control-group estimator has an obvious attraction--estimated program effects require only the calculation of mean changes in macroeconomic outcomes for program and for nonprogram countries, i.e., only $(\overline{\Delta y_j})_p$ and $(\overline{\Delta y_j})_N$. The estimation requirements for the modified control-group estimator, however, are substantially more demanding. Not only do we need values for three additional variables or parameters (namely, $\bar{x}_N$, $\bar{x}_p$ and β_j), but we also face the problem that two of them, $\bar{x}_p$ and β_j, are not observed directly. Recall that $\bar{x}_p$ is not observed because x_i refers to policies that would have been undertaken in the absence of programs; thus, x_i is equal to observed policies in non-program countries but not in program countries. Hence, implementation of the modified control-group approach requires estimating x_i for program countries (as well as the estimation of the parameter β_j linking Δx_i to Δy_{ij}).

The first step in estimating x_i for program countries is to fit the reaction-function equation (2) to observable data for non-program countries. The only unobserved variable in equation (2) is the country-specific vector of desired macroeconomic outcomes, y_i^d. If this variable can be assumed to be constant over time, it can be captured by a set of country-specific constants, giving the policy-reaction equation the following final form:

$$(22) \quad \Delta x_i = \gamma_{0i} - \gamma(y_i)_{-1} + \eta_i$$

The fitted values of this equation for program countries constitute the counter-factual Δx_i. In effect, this procedure uses data on observed policy behavior in non-program countries to identify "normal" policy-reaction in given policy-target circumstances. 1/ This normal policy reaction is then used to estimate what "would have been" in program countries if there had not been a Fund program.

An important caveat is in order here concerning another source of potential systematic differences between program and non-program countries. Since both the setting of policy instruments in equation (22) and the acceptance by a country of a Fund program as specified in equation (3) reflect policy decisions of the authorities, any unobservable factors (π_i) that make a given country more likely to go to the Fund--such as a general "committment to adjustment"-may also make that country more likely to have adopted a different policy package in the absence of a program, Δx_i, than another country facing similar observable (policy-target) circumstances. In this case, the behavior of non-program countries would not be a good guide to the "counter-factual" in program countries--even after observable pre-program characteristics of the two groups are controlled for. Formally, this possibility would manifest itself in our model via correlation between the error term π_i in equation (3) and η_i in (2). If such a correlation is present, then equation (22) will provide a biased estimate of Δx_i for program countries--essentially because it fails to remove this aforementioned additional source of sample-selectivity bias. 2/

1/ In a pooled cross-section time-series sample, this would include observations in non-program periods for countries which are program countries in other periods.

2/ The direction of the bias depends in part on the correlation between π_i and η_i. Intuitively, if for a given set of observable pre-program circumstances, countries that would have pursued "worse" policies are more likely to adopt Fund programs, then the behavior of non-program countries would provide an excessively optimistic counter-factual and the beneficial effects of programs would be understated. Conversely, if programs are more likely to be adopted by countries that would have undertaken "better" policies anyway, then the beneficial effects of programs would be overstated because the favorable effects of the policies would be erroneously attributed to the fact of Fund involvement.

This additional source of bias can be eliminated even though both η_i and π_i are unobservable. The reader is referred to Heckman [1979] for a description of the appropriate procedure. For our purposes, we note that the procedure requires the specification and estimation of a model of program participation--i.e., of equation (3). Thus, removal of the two sources of sample selectivity bias we have identified requires the specification and estimation of models of endogenous policy formation (equation (2)) and program participation (equation (3)).

With $(\bar{y}_j)_P$, $(\bar{y}_j)_N$, and $\bar{x}_N$ observed directly and with $\bar{x}_P$ estimated as outlined above, the remaining element necessary for application of the modified control-group estimator is the parameter vector β_j, which links normal policy changes in the absence of programs to changes in the macroeconomic target variables. Up to now, this vector has been assumed to be known. For our purposes, any unbiased estimator of β_j will do the job. Perhaps the simplest way to produce such an estimate is to fit the macroeconomic outcome equation (20) in level form to a pooled cross-section time-series data sample by using observed values for the policy vector x_i. 1/

If the objective is solely to obtain an unbiased estimate of total program effects, we can substitute the policy-reaction equation (2) for Δx_i into the level-form equation (20) and derive:

$$(23)\quad y_{ij} = \beta_{0i} - (y_{ij})'_{-1}\,\gamma'\beta_j - (x'_i)_{-1}\,\beta_j + W'\alpha_j + \beta_j^{IMF}\,d_i + (\varepsilon_{ij} + \eta_i\beta_j)$$

Fitting equation (23) to observable data will then yield an estimate of total program effects via the estimated coefficient (β_j^{IMF}) on the dummy variable d_i. However, this does not take into account any sample-selectivity bias arising from systematic differences in reaction functions as between program and non-program countries. If the error terms in equation (2) and (3) are in fact correlated, then the reduced-form approach has to be augmented by the Heckman [1979] correction in order to obtain unbiased estimates of program effects. This short-cut works because it essentially controls for observable differences between program and nonprogram countries. However, it cannot yield information on how total program effects are apportioned as between changes in policy instruments and other factors.

Analyzing how programs work

To analyze the three different channels by which programs can affect macroeconomic outcomes, it is helpful to introduce some additional notation. Let $x_{i,IMF}$ be the vector of policy instruments adopted

1/ If Fund programs induce parameter shifts, then only data on nonprogram countries could be used for this purpose.

under a program, $\beta_{j,IMF}$ the vector of coefficients linking these policy instruments to the target variable y_{ij}, and $CON_{i,IMF}$ any unmeasurable "confidence" effects on y_{ij} attributable to a program. As before, x_i and β_j will be the values of policy instruments and their coefficients in the absence of a program. We can then express the total effect of a Fund program (β_j^{IMF}) as:

$$(24) \quad \beta_j^{IMF} = (x_{i,IMF}' \beta_{j,IMF} - x_i' \beta_j) + CON_{i,IMF}$$

Rewriting the level-form equation (20) for y_{ij} with the substitution for β_j^{IMF} yields:

$$(25) \quad y_{ij} = x_i' \beta_j + W'\alpha_j + [CON_{i,IMF} + (x_{i,IMF}' \beta_{j,IMF} - x_i' \beta_j)] d_i + \varepsilon_{ij}$$

It is clear that if separate estimates of $\beta_{j,IMF}$ and $CON_{i,IMF}$ could be obtained, it would be possible to identify the separate channels through which a program affects y_{ij}. Given the estimate of x_i for program countries and the estimate of β_j as outlined above, we next need to estimate the following equation:

$$(26) \quad y_{ij} = x_i'(1-d_i) \beta_j + x_{i,IMF}' d_i \beta_{j,IMF} + CON_{i,IMF} d_i + \varepsilon_{ij}$$

The estimated coefficients on $x_{i,IMF} d_i$ and on d_i will then be the estimates of $\beta_{j,IMF}$ and of $CON_{i,IMF}$ that we seek. Of course, if estimation of (26) produces the result that β_j is not significantly different from $\beta_{j,IMF}$, then we can put aside shifts in behavioral parameters as a source of program effects and deal exclusively with $(x_{i,IMF}-x_i)'\beta_j$ and $CON_{i,IMF}$.

To summarize, in this section we have shown that the presence of systematic differences between program and non-program countries need not render the control-group approach to the estimation of program effects useless. One way of handling the problem is to account for any differences in indigenous macroeconomic policy between program and non-program countries and to employ the level of macroeconomic performance in the program period rather than its change. However, this "modified" estimator is significantly more difficult to calculate than the traditional control-group estimator. On the other hand, one important feature of the more structural version of this estimator is that it can be used to provide information not only on total program effects but also on <u>how</u> these effects are apportioned as among induced changes

in policy instruments, shifts in behavioral parameters, and general "confidence" effects. 1/

IV. Some Empirical Exercises

Demonstrating that several alternative estimation methods can in theory yield different results about the effects of Fund progams is one thing. Illustrating the empirical relevance of that point with actual data on Fund programs is quite another. In this section we provide an exploratory empirical investigation of the aforementioned methodological pitfalls by comparing estimates of Fund program effects across the three estimators discussed earlier, namely a before-after comparison of mean outcomes for programs countries alone (estimator A), a before-after comparison for program countries relative to that for non-program countries (the traditional control-group estimator B), and a reduced-form regression estimate of program effects that controls only for observed pre-program differences between program and non-program countries (a version of the modified control-group estimator M). The data samples are drawn from the population of Fund stabilization programs over the 1974-81 period.

As suggested earlier, while we think these empirical results are instructive for testing the sensitivity of estimated program effects to alternative estimation methods, we do not think that much confidence ought to be placed in any of the estimates of program effects themselves. We say this because the particular equations tested, even for the modified control-group estimator, accomodate only one of the possible sources of bias outlined in sections II-III (we have not investigated the empirical relevance of correlations between the unobservable components of policy-reaction functions and the factors affecting program participation); because we do not construct a carefully-specified, structural, economic model for the macroeconomic-outcomes variables or for indigenous policy reaction; because we have experimented with only one short time span for program effects (i.e., the change from the pre-program year to the program year) 2/; and because the goodness-of-fit characteristics of the

---

1/ It should be acknowledged that if economic and political conditions change markedly at frequent intervals and/or if governments with different policy-reactions functions appear frequently, then it may not be feasible to empirically identify a "stable" policy-reaction function. But this is a matter for empirical testing.

2/ The question of when a program-country stops being a program country is a particularly difficult one to answer, yet it can have an important effect on program estimates based on multi-country data. Suppose, for example, that two countries face identical current account deficits. Country A, with a Fund program, undertakes a devaluation cum expenditure reducing policy while country B without a program adopts increased trade restrictions. Over a one year period, the change in the current account could well be quite similar for the two

estimates themselves do not merit such confidence. Having said that, we should also point out that most of the same deficiencies also plague (to our knowledge) all the earlier empirical literature on program effects using multi-country data. 1/

The data base

Our estimates were made using a sample which contains observations from 58 developing countries during 1974-81 sample period. It consists of 397 country-year observations, 68 of which are program-year observations. The 58 countries in the sample are those for which data were available for all relevant macroeconomic variables for at least two consecutive years during 1974-81 (not necessarily for the entire period). Consecutive-year Fund programs, including those classified as Extended Fund Facility programs, are included. A list of the program countries included is presented in Appendix Table A1. 2/

Definition of variables

As in most earlier studies, we have selected some popular indicators of external and internal balance as the appropriate "outcome" or "target" variables for Fund stabilization programs. Specifically, the four outcome variables that serve as the empirical counterparts to the y variables of the theoretical sections are: the ratio of the overall balance of payments to nominal gross national product (BOP/GNP), the ratio of the current account of the balance of payments to nominal gross national product (CA/GNP), the rate of inflation as measured by the consumer price index ($\Delta CPI/CPI_{t-1}$), and the rate of growth of real gross domestic product ($\Delta RGDP/RGDP_{t-1}$). These four summary indicators are of course not the only relevant yardsticks of the success of a Fund

---

1/ (Cont'd from p. 25) countries. Over a longer period (after the program), one might expect country A to show better growth and external balance performance than country B. But this would not be reflected in one-year comparisons. Indeed, country A would be classified as a non-program country after the program year.

1/ In this respect, the attention devoted by Donovan [1982] to both long and short-term effects of programs, and by Gylfason [1983] to the theoretical channels by which domestic credit can affect economic growth as well as the balance of payments, are particularly commendable.

2/ We also ran some tests on several smaller samples. Since the results were qualitatively similar to those reported here, we did not include them in the text.

program but it would be hard to argue that they are not important ones. 1/ Also, for the purposes of this study, they carry the advantage of facilitating comparison with earlier empirical work on program effects. 2/

Recall from Section III that calculation of the modified control-group estimator requires data on the vector of policy instruments for both program and non-program countries. For this purpose, we collected data on total domestic credit (D) and on the real effective exchange rate (REX) for each of our sample countries. 3/ They serve as the empirical counterpart to the x variables of the theoretical sections. Again, it is not difficult to think of other policy instruments that would be pertinent to Fund stabilization programs but few would deny the key roles accorded to these two instruments in most programs.

Finally, in order to create the dummy variable d_i that captures the presence (d_i=1) or absence (d_i=0) of a Fund program, we assigned a program to a given year if it was approved (by the Fund's Executive Board) during the first six months of that year. Otherwise, it was assigned to the following year. Also, the phrase "program countries" is used in what follows to refer to those (country-year) observations during which Fund programs were in effect.

A listing of the data sources for each of our variables is presented in Appendix Table A2.

How the estimators were calculated

All that remains before examining the results themselves is to briefly review how the three alternative estimators of Fund program effects were actually calculated in the tables that follow.

For the simple before-after estimator ($\hat{\beta}_j^{IMF,A}$), we computed the mean change across our group of program countries for each of the four macroeconomic outcome variables. In terms of earlier symbols, the before-after estimator is then:

$$(33) \quad \hat{\beta}_j^{IMF,A} = (\overline{\Delta y_j})_P$$

1/ In addition to the four indicators mentioned above (measured somewhat differently, Donovan [1982] also examined changes in savings and investment ratios and changes in the growth rate of real consumption.

2/ In some of the earlier studies, the external balance variables were scaled by nominal exports rather than by nominal GNP but we doubt that this difference has any material effect on the qualitative nature of the results.

3/ The real effective exchange rate is an import-weighted index with the CPI employed as the relevant deflator; see Appendix Table 3 for a more precise definition.

For computational convenience, the traditional control-group estimator was calculated by running the following regression equation on the combined sample of program and non-program countries:

$$(34) \quad \Delta y_i = \alpha_1 + \alpha_2 d_i$$

where (recall) d_i is the dummy variable for Fund program status and where α_1 and α_2 are estimated coefficients. The estimate of α_2 will then be our traditional control-group estimator, $\hat{\beta}_j^{IMF,B}$.

Last, we have the modified control-group estimator $\hat{\beta}_j^{IMF,M}$. As suggested in Section III, there are several ways to calculate it. Here, since our primary purpose is to determine how sensitive estimated program effects are to alternative assumptions, it seemed acceptable to concern ourselves only with total program effects. We therefore chose to use the reduced-form version of the modified control-group estimator given in equation (23) since it is so much easier to calculate; again, we did not correct for any possible correlation between the unobservable components of program participation and of the policy-reaction function.

By subtracting $(y_j)_{-1}$ from both sides of (23), this equation can be estimated in the form:

$$(23') \quad \Delta y_{ij} = \beta_{0i} - \sum_h (y_{ih})_{-1} \lambda_h - (1+\lambda_j)(y_{ij})_{-1} - (x_i')_{-1} \beta_j + W' \alpha_j + \beta_j^{IMF} d_i + (\varepsilon_{ij} + \eta_i' \beta_j)$$

where $\lambda = \gamma' \beta_j$ is an N X 1 vector with jth element equal to λ_j. As a proxy for W, the variable which measures the international economic environment, we introduced a set of time dummy variables. Also, the β_{0i} are coefficients of a set of country dummies designed to capture inter-country differences in desired target values for the y_{ij}. It is also possible to test formally whether the additional variables peculiar to the modified control-group estimator, namely the lagged values of the vectors y_i and x_i, make as a set a significant contribution to the explanation of Δy. To do so, one performs an F-test on the null hypothesis that the coefficients of these variables are all zero. Observe also that even if prior statistical tests document that program countries differ systematically from non-program countries with respect to these variables, these pre-program period characteristics must show a statistically significant effect as a group on Δy for there to be a bias in the traditional control-group estimator of program effects. For if this latter condition is not satisfied, then the

estimates of program effects using the traditional and modified control-group methodologies will yield the same result. 1/

The results

The results of main interest are laid out in Tables 1-5. Tables 1 and 2 provide estimates of program effects under the before-after estimator and the traditional control-group estimator, respectively. Table 3 presents the results of a test for differences between program and non-program countries in the level of macroeconomic outcomes prior to the program period. Table 4 gives the estimates of program effects using the modified control-group estimator. Finally, Table 5 presents a capsule summary of the sensitivity of estimated program effects to the estimation methodology.

Table 1, although it is confined to changes in macroeconomic outcomes for program countries alone, already raises some doubts about the quality of estimates based on simple before-after calculations. There is a marked difference in the nature and pattern of estimated program effects from year to year. Note, for example, the difference in estimated program effects as between say, 1976 programs and 1980 programs. Again, while it is possible that "true" program effects really do change markedly from year to year, it seems more likely that this temporal instability arises because the non-program determinants of changes in macroeconomic outcomes (e.g., oil shocks, foreign demand conditions, agricultural supply shocks, etc.), often change significantly from year to year. Since the before-after methodology does not acknowledge such non-program influences on Δy, it cannot control for them in estimating program effects.

Turning to Table 2, which conveys the traditional control-group estimates of program effects, there are three features of the results worth mentioning. First, the sizes and even the direction of estimated

---

1/ Some readers will recognize this as an application of "specification bias due to an omitted variable"; see, for example, Kmenta [1971], p. 391. In brief, if equation (23) is the "true model" but we estimate equation (34) instead, then the bias attaching to α_2 in equation (34) will be a product of two factors: (i) the correlations between the program dummy d_i and the omitted variables (here, the lagged values of y and x); and (ii) the coefficients on the omitted variables. If either (i) or (ii) is zero, then the α_2 in equation (34) will equal the β_j^{IMF} in equation (23).

Table 1

Before-After Estimates of Program Effects for Individual Years 1/

(In percent)

	1974	1975	1976	1977	1978	1979	1980	1981	1974-81
Δ(BOP/GNP)	-0.05	-2.78	2.84	1.59	-4.35	4.00	-1.69	-1.69	-0.37
Δ(CA/GNP)	-4.40	-2.41	4.50	2.61	-1.26	-0.69	-1.77	-1.12	-0.28
Δ(ΔCPI/CPI)	52.45	-32.33	-13.93	-22.90	5.39	5.15	2.61	-1.45	-3.05
Δ(ΔRGDP/RGDP)	0.02	- 0.07	0.05	- 0.00	0.00	0.00	-0.02	0.01	-0.09

1/ Variables are defined on page 26 of the text.

Table 2. Traditional Control Group Estimates of Program Effects

(In percent)

Mean change in outcome variables	Program countries	Non-program countries	Difference	t-statistic
Δ (BOP/GNP)	-0.37	-0.21	-0.16	-0.19
Δ (CA/GNP)	-0.28	-0.72	0.44	0.51
Δ (ΔCPI/CPI$_{t-1}$)	-3.05	0.95	-4.00	-1.06
Δ (ΔRGDP/RGDP$_{t-1}$)	-0.09	-0.25	0.16	0.20

program effects sometimes change quite noticeably from those obtained under the before-after estimator. Specifically, once the performance of non-program countries is used as a measuring rod, Fund programs now are associated with a reduction in inflation and slightly better growth. Second, Table 2 documents the importance of applying tests of statistical significance to observed differences in performance between program and non-program countries. Whereas the macroeconomic performance of program countries is always different from that of non-program countries in each of the four comparisons shown in Table 2, in none of them could it be legitimately concluded that the observed difference was statistically significant (i.e., not the outcome of chance).

As emphasized in the preceding sections, we must suspect that the traditonal control-group estimates of program effects will be biased if the selection of program countries is non-random and if these non-random characteristics are correlated with macroeconomic performance during the program period. Tables 3 and 4 address those two questions. In particular, Table 3 tests our earlier argument that Fund program status is likely to be systematically related to the country's level of macroeconomic performance prior to the program period. The results are straightforward and can be summarized as follows. Program countries do seem to be different than non-program countries. In the year before the inception of a Fund program, program countries experienced (on average) larger balance of payments deficits in proportion to GNP, larger current account deficits in proportion to GNP, higher rates of inflation, and lower rates of real output growth than did non-program countries. Each of these differences is statistically significant at the 5 percent level or better. This is revealed not only by the t-test results shown in Table 3 but also by chi-squared tests for differences in the whole set of mean comparisons. These differences in pre-program conditions as between program and non-program countries appear in all samples we examined. Indeed, the existence of these pre-program differences between program and non-program countries was the single most "robust" empirical finding in all our tests.

Table 4 takes the case one step further by testing whether these revealed pre-program differences in macroeconomic outcomes affect the change in macroeconomic performance as between the pre-program year and the year of the program. Again, the results of interest can be conveniently summarized. First, pre-program levels of macroeconomic outcomes do appear to affect the change in these outcomes. In all four equations shown in Table 4, the change in the outcome variable is related in a statistically significant way to two or more of the four outcome-level variables in the pre-program year. In each case, the outcome-change variable is related to its own lagged level and

Table 3. Differences Between Program and Non-Program Countries: Means of Outcome Variables in Pre-Program Year

(In percent)

Outcome variables	Program countries	Non-program countries	Difference	t-statistic
BOP/GNP	-1.53	1.43	-2.96	-3.58**
CA/GNP	-7.26	-5.26	-2.00	-2.10**
$\Delta CPI/CPI_{t-1}$	41.51	20.44	21.07	3.36**
$\Delta RGDP/RGDP_{t-1}$	3.82	5.23	-1.41	-2.02**

** Indicates statistical significance at the 1 percent level.

Table 4. Modified Control-Group Estimates of Program Effects [1/]

Dependent Variable	$(BOP/GNP)_{t-1}$	$(CA/GNP)_{t-1}$	$(\Delta CPI/CPI)_{t-1}$	$(\Delta RGNP/RGNP)_{t-1}$	$(\Delta D/D)_{t-1}$	REX_{t-1}	d_i
$\Delta(BOP/GNP)$	- 1.092**	0.056	-0.017*	- 0.001	0.001	-00.023	-0.011
	(0.063)	(0.055)	(0.009)	(0.001)	(0.002)	(00.019)	(0.008)
$\Delta(CA/GNP)$	- 0.141*	- 0.702**	0.000	- 0.002**	0.001	00.039	-0.001
	(0.072)	(0.063)	(0.000)	(0.001)	(0.003)	(00.022)	(0.009)
$\Delta(\Delta CPI/CPI_{t-1})$	-71.990*	32.117	-0.458**	-0.351	0.012	-19.433	3.248
	(35.976)	(31.665)	(0.049)	(0.310)	(0.013)	(10.877)	(4.448)
$\Delta(\Delta RGDP/RGDP_{t-1})$	27.504**	- 6.192	-0.010	-0.969**	-0.001	- 3.179	-0.220
	(5.889)	(5.183)	(0.008)	(0.051)	(0.002)	(1.780)	(0.728)

Figures in parentheses are standard errors.
Coefficients of time and country-specific dummy variables are not reported.

* Indicates statistical significance of the 5 percent level.
** Indicates statistical significance at the 1 percent level.

[1/] The variables BOP/GNP and CA/GNP are measured as fractions, whereas $\Delta CPI/CPI_{t-1}$ and $\Delta RGDP/RGDP_{t-1}$ are in percent.

with a negative coefficient. This can be taken as supportive of the notion advanced earlier that macroeconomic outcomes in both program and non-program countries may display a "regression to the mean" property. For example, the greater the size of the current account deficit in period t-1, the greater the improvement in the current account as between period t-1 and period t. Second and not surprising, estimated program effects under the modified control-group estimator are quite different from those obtained under the traditional control-group estimator. This can be seen most vividly in Table 5 where the three estimators are shown side-by-side. Also, and consistent with our earlier expectations about the direction of bias, we find that estimated program effects, after allowing for pre-program differences between program and nonprogram countries, are almost always less favorable. For example, the improvement in the current account ratio disappears entirely, the deterioration in the BOP ratio is magnified, and the favorable outcomes for inflation and growth are reversed. Third, employing tests of statistical significance again indicates that observed differences in macroeconomic performance between program and non-program countries are not significant. Fourth and finally, while the explanatory power of the regression equations in Table 4 is rather low (in the range of .1 to .2 without country and time dummy variables and .2 to .3 with them), the explanatory power is significantly higher (in a statistical sense) when those variables peculiar to the modified control-group estimator are included in the equations. In this respect, an F-test reveals that for each of the equations given in Table 4, the modified control-group variables (i.e., the lagged values of targets and instruments) are statistically significant as a group at the 1 per cent level. Put in other words, the modified control-group equations in Table 4 hardly provide a "full" or even a "good" explanation for observed changes in macroeconomic outcomes, but they are significantly better than would be obtained by the implied equations of the traditional control-group approach.

To summarize, we have shown in this section, if only in a preliminary way, that it does make a significant difference how one estimates the effect of programs from cross-sectional data. Thus, some of the theoretical sources of bias outlined in the earlier sections do appear to be of more than academic interest.

VI. Conclusions

Given the pivotal role assigned to Fund stabilization programs in the past and present economic policy strategies of so many developing countries, and given the continuing controversy on their effects, it is not suprising that there has been strong interest in empirical measures of program effectiveness. Also, because the large number of

Table 5. Comparisons of Alternative Estimators of Program Effects

(In percent)

Outcome Variable	Before-After Approach	Traditional Control-Group Approach	Modified Control-Group Approach
Δ(BOP/GNP)	-0.37	-0.16	-1.10
Δ(CA/GNP)	-0.28	0.44	-0.10
$\Delta(\Delta CPI/CPI_{t-1})$	-3.06	-4.00	3.25
$\Delta(\Delta RGDP/RGDP_{t-1})$	-0.09	0.16	-0.22

such programs makes the "case-by-case" approach such a laborious and time-consuming way to arrive at an estimate of "average" program effectiveness, it is likewise understandable that the cross-country approach to program evaluation has dominated the empirical literature. We have argued in this paper, however, that if the estimated program effects from such a cross-country analysis are to be representative of "true" program effects, then certain methodological pitfalls need to be avoided. At the risk of ignoring some problems and of unduly simplifying others, the main lessons of the preceding analysis can be summarized as follows:

(1) A before-after comparison of mean macroeconomic outcomes for program countries is unlikely to yield a good estimate of true program effects because the non-program determinants of macroeconomic outcomes typically change as between the pre-program period and the program period. As such, ascribing all of the observed change in outcomes to the program alone will invariably overstate or understate the true independent effect of the program.

(2) If the mean change in outcomes for non-program countries is subtracted from the mean change for program countries, the bias in program estimates due to ignoring the non-program determinants of macroeconomic outcomes will be reduced. However, a new source of bias will be introduced whenever program countries differ systematically from non-program countries in some characteristic that is related to subsequent macroeconomic performance. In the particular case when the determinants of Fund program selection are positively correlated with the non-program determinants of changes in macroeconomic outcomes, this traditional control-group estimate of program effects will overstate true program effects. Further, preliminary empirical tests suggest that in practice (at least for the 1974-81 period) program countries did have significantly less favorable macroeconomic peformance than non-program countries prior to the program period, and that such pre-program outcomes were significantly related to subsequent performance during the program period itself. Not surprisingly therefore, estimates of program effects that held constant the pre-program levels of macroeconomic outcomes were quite different than those that did not. In any case, the moral is that if the program countries are not selected randomly, then these non-random selection criteria must be identified so that either a control group can be found with the same characteristics or so that these group differences can be accounted for in any comparison of outcomes as between the two groups.

(3) Because Fund programs probably work in good measure by changing the stance of policy instruments from what it would be in the absence of a program, any estimate of program effects that does not allow for this channel of influence runs the risk of capturing only part of total program effects (e.g., just "confidence" effects) and thus of

understating true program effectiveness. 1/ In this paper, we have outlined an estimation procedure which in principle permits a calculation of how total program effects are apportioned as among induced changes in policy instruments, induced changes in behavioral parameters, and general confidence effects. Central to this procedure is the estimation of "policy reaction functions" for both program and non-program countries. While we would not want to underestimate the practical difficulties associated with obtaining credible estimates of such reaction functions for developing countries (particularly when underlying economic and political conditions are changing markedly at frequent intervals), we see no other way of estimating the "counter-factual" for program countries. Also, information on how programs work may be just as important for program design as information on total program effects.

(4) In comparing the performance of program countries to that of non-program countries, it is strongly advisable to subject any differences to tests of statistical significance. As brought out in our empirical investigation, it frequently turns out that observed differences in performance between the two groups during the program period would not be judged statistically significant at conventional confidence levels.

(5) On a broader level, the methodological problems we have described lead us to the view that considerable caution is needed in attempting to estimate and interpret the effects of Fund programs using multi-country data.

1/ For an example of such an estimate, see Chapman and Killick [1982].

APPENDIX

Bias in the Traditional Control-Group Approach

Equations (14a) and (14b) in the text are crucial for establishing the presence of bias in the traditional control-group approach under non-random selection of program countries. In the first part of the Appendix, we derive these equations.

Denote the variance of $(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})$ as σ_x^2. Suppose that $\Delta x_i' \beta_j + \Delta \varepsilon_{ij}$ and z_i have a joint normal distribution with the correlation between $\Delta x_i' \beta_j + \Delta \varepsilon_{ij}$ and z_i denoted as ρ_{xz}. Finally, let ϕ and Φ represent the standard normal density and distribution functions respectively. For the ith country, it will then be true that:

$$(A1)\quad E[z_i/(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})] = \rho_{xz} \frac{\sigma_z}{\sigma_x} (\Delta x_i' \beta_j + \Delta \varepsilon_{ij} - \Delta x' \beta_j)$$

The probability (Pr) that country i will be a program country is then:

$$(A2)\quad Pr(i \varepsilon P) = Pr(z_i > z^*) = 1 - \Phi[(z^* - \rho_{xz} \frac{\sigma_z}{\sigma_x} (\Delta x_i \beta_j + \Delta \varepsilon_{ij} - \Delta x' \beta_j))/\sigma_z]$$

In the special case where ρ_{xz} is zero, equation (A2) reduces to

$$(A2')\quad Pr(i \varepsilon P) = 1 - \phi[z^*/\sigma_z]$$

The key difference between (A2) and (A2') is that while the probability of being a program country in (A2) is a function of $\Delta x_i' \beta_j + \Delta \varepsilon_{ij}$ and thus will differ across countries, in (A2') this probability is the same for all countries in the sample.

The next step to discovering the direction of bias in the control group estimate of program effects under conditions of non-random program selection is to write:

$$(A3)\quad E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/z_i > z^*] = \Delta x' \beta_j + \rho_{xz} \frac{\sigma_x \phi(z^*/\sigma_z)}{1 - \Phi(z^*/\sigma z)}$$

$$(A4)\quad E[(\Delta x_i' \beta_j + \Delta \varepsilon_{ij})/z_i \leq z^*] = \Delta x' \beta_j - \rho_{xz} \frac{\sigma_x \phi(z^*/\sigma_z)}{\Phi(z^*/\sigma_z)}$$

APPENDIX

Since σ_x and ϕ are both positive, and since Φ is bounded between zero and one, $\sigma_x\phi/(1-\Phi)$ and $\sigma_x\phi/\Phi$ are both positive. Equations (14a) and (14b) in the text then follow directly from (A3) and (A4) respectively.

APPENDIX

Table A1.

The sample contains the program countries listed below.

1974

Chile, Pakistan, Sri Lanka

1975

Burma, Chile, Israel, Pakistan

1976

Bangladesh, Haiti, Kenya, Korea, Philippines, Tanzania

1977

Argentina, Burma, Egypt, Haiti, Israel, Kenya, Mexico, Pakistan, Philippines, South Africa, Zaire, Zambia

1978

Argentina, Jamaica, Kenya, Mexico, Peru Philippines, Portugal, Sri Lanka, Turkey, Zambia

1979

Burma, Egypt, Haiti, Honduras, Jamaica, Mexico, Peru, Philippines, Sri Lanka

1980

Bangladesh, Bolivia, Haiti, Honduras, Jamaica, Kenya, Korea, Malawi, Peru, Philippines, Sierra Leone, Sri Lanka, Turkey, Zaire

1981

Bangladesh, Burma, Haiti, Honduras, Jamaica, Korea, Pakistan, Sierra Leone, Thailand, Zambia.

APPENDIX

Table A2.

The sources of the data used in Section V are as follows:

a) Net foreign assets, International Financial Statistics, line 31.

b) Nominal GNP - Current Studies division data file.

c) Real GDP - Current Studies division data file.

d) Current account - Current Studies division data file.

e) Consumer price index - International Financial Statistics, line 64.

f) Domestic Credit - International Financial Statistics, line 32.

g) Real effective exchange rate - Developing Countries Studies division data file. The precise definition is:

$$REX \equiv 100.e^{\left[\sum_{s=1}^{n} \ln(EXI_i/EXI_s)W_s - \sum_{s=1}^{n} \ln(CPI_s/CPI_i)W_s\right]}$$

EXI = nominal exchange rate index

i = reporting country

s = partner country

W_s = import weight for partner country s.

CPI = consumer price index

References

Ashenfelter, O., "Estimating the Effect of Training Programs on Earnings," Review of Economics and Statistics, February 1978, pp. 47-57.

_____ and D. Card, "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," Working Paper No. 174, Industrial Relations Section, Princeton University, July 1984.

Beveridge, W.A., "Fiscal Adjustment in Financial Programs Supported by Stand-By Arrangements in the Upper Credit Tranches, 1978-79," (unpublished, International Monetary Fund, July 1, 1981).

_____, and M. Kelley, "Fiscal Content of Financial Programs Supported by Stand-By Arrangements in the Upper Credit Tranches," IMF Staff Papers, June 1980, pp. 205-49.

Chapman, M., and T. Killick, "Much Ado About Nothing? Testing the Impact of IMF Stabilization Programmes in Developing Countries," Overseas Development Institute Working Paper No. , 1982.

Connors, T., "The Apparent Effects of Recent IMF Stablization Programs," Federal Reserve International Finance Discussion Paper No. 135, April 1979.

Cline, W. and S. Weintraub, eds., "Economic Stablization in Developing Countries, Washington, D.C., 1981.

Donovan, D., "Real Responses Associated with Exchange Rate Action in Selected Supper 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.

Guitian, M., Fund Conditionality: Evolution of Principles and Practices, International Monetary Fund Pamphlet Series, No. 38, 1981.

Goldstein, M., and Mohsin Khan, Effects of Slowdown in Industrial Countries on Growth in Non-Oil Developing Countries, Occasional Paper No. 12, International Monetary Fund, August 1982.

_____, The Global Effects of Fund-Supported Adjustment Programs, Occasional Paper No. 42, International Monetary Fund, forthcoming, 1986.

Gylfason, T., "Credit Policy and Economic Activity in Developing Countries: An Evaluation of Stand-By Programs, 1977-79," (unpublished, Institute for International Economic Studies, University of Stockholm, 1983).

Heckman, J., "Sample Selection Bias as a Specification Error", Econometrica, January 1979, pp. 153-60.

International Monetary Fund, World Economic Outlook, Washington, D.C., 1983.

Kelly, M., "Fiscal Adjustment and Fund-Supported Programs, 1971-1980" (unpublished, International Monetary Fund, September 28, 1982).

Khan, M., and M. Knight, "Determinants of the Current-Account Balances of Non-Oil Developing Countries in the 1970s: An Empirical Analysis," (unpublished, International Monetary Fund, June 28, 1983).

Killick, T., The Quest for Economic Stabilization: The IMF and the Third World, St. Martin's Press, New York, 1984.

Kmenta, J., Elements of Econometrics, Macmillan Press, New York, 1971.

Lewis, H.G., Unionism and Relative Wages in the United States: An Empirical Inquiry, University of Chicago Press, 1963.

Loxley, J., The IMF and the Poorest Countries, The North-South Institute, 1985.

Lucas, R. E., Jr., "Econometric Policy Evaluation: A Critique," Journal of Monetary Economics, Vol. 2, Supplement, 1976, pp. 19-46.

Nowzad, B., The IMF and its Critics, Essays in International Finance No. 146, International Finance Section, Department of Economics, Princeton University, December 1981.

Odling-Smee, J., "Adjustment with Financial Assistance from the Fund: The Experience of Seven Countries," Finance and Development, December 1982, pp. 26-30.

Reichmann, T., "The Fund's Conditional Assistance and the Problems of Adjustment, 1973-75," Finance and Development, December 1978, pp. 38-41.

_____, and R. Stillson, "Experience with Balance of Payments Adjustment: Stand-By Arrangements in the Higher Credit Tranches, 1963-72," IMF Staff Papers, June 1978, pp. 293-309.

Williamson, J., The Lending Policies of the International Monetary Fund, Institute for International Economics, Washington, D.C., August 1982.

_____ (ed)., IMF Conditionality, MIT Press, Cambridge, Mass., 1983.

Zulu, J., and S. Nsouli, Adjustment Programs in Africa: The Recent Experience, IMF Occasional Paper No. 34, Washington, D. C., 1985.