IMF WORKING PAPER

WP/95/92

INTERNATIONAL MONETARY FUND

The Macroeconomic Effects of ESAF-Supported Programs: Revisiting Some Methodological Issues

Prepared by

Louis Dicks-Mireaux, Mauro Mecagni and Susan Schadler 1/

September 1995

Abstract

This paper examines whether ESAF-supported programs during 1986-91 had significant independent effects on growth, inflation and the external debt service ratio. Econometric estimates of the Generalized Evaluation Estimator (GEE) identify statistically significant beneficial effects on output growth and the debt service ratio but no effects on inflation. The robustness of these estimates is also examined. Diagnostic tests cast doubt on the applicability of the GEE framework to the ESAF-eligible countries, and the results obtained using it.

JEL Classification Numbers:
E65; F33

1/ The authors are indebted to Sharmini Coorey for her considerable input during the early stages of the project and to Kirsten Fitchett for her invaluable research assistance. We also thank Adam Bennett, David Burton, Benedicte Vibe Christensen, David Goldsbrough, Morris Goldstein, Mohsin Khan, Peter Montiel, Roger Nord, and Alessandro Zanello for their comments on an earlier draft of this paper. The views expressed in this paper are those of the authors and do not necessarily represent those of the International Monetary Fund.

Contents | Page

Contents | Page

Summary

A number of different questions and methodologies can be used to evaluate Fund-supported programs. This paper seeks to answer one of these questions: what are the independent effects of ESAF-supported programs on growth, inflation, and the external debt service ratio, relative to what would have been in the absence of an ESAF-supported arrangement? To do so, the Generalized Evaluation Estimator (GEE) is applied to ESAF-supported programs during 1986-91. Within the framework of the GEE, a counterfactual of policies that would have been implemented in the absence of Fund support is constructed and the effects of counterfactual policies, exogenous developments, initial conditions, and the involvement of the Fund on macroeconomic performance are econometrically estimated. These estimates of the GEE, and a number of elaborations aimed at broadening its application, identify statistically significant beneficial independent effects of ESAF support on output growth and the debt service ratio. The effects on inflation were not statistically significant from zero.

This study also considers whether, for the sample of ESAF-eligible countries, the assumptions required for the GEE to yield reliable estimates of the independent effects of Fund-supported programs are satisfied. The GEE relies on two basic presumptions: that a relatively simple macroeconomic model can reasonably capture the interaction between macroeconomic policies and outcomes for a large number of countries over time; and that counterfactual policies can be characterized by a policy reaction function estimated for a period when, and in countries where, Fund support is not in place. Diagnostic tests cast doubt on the applicability of the GEE framework to the ESAF-eligible countries. In particular, the estimates are sensitive to variation in the sample, and a counterfactual policy reaction function with significant explanatory power could not be identified.

Most recent applications of the GEE contain little or no evaluation of the validity of the underlying model. One important lesson of this study is that the validity for any given sample of the premises of the GEE methodology must be investigated before reliable conclusions about the independent effects of Fund supported programs can be drawn from it. The paper also points to some elaborations of the basic GEE framework that would likely help reduce its inherent restrictiveness and perhaps make it applicable in a broader range of circumstances. However, substantially more experimentation will be needed to produce dependable estimates of the independent effects of Fund support. Such experimentation will need to address several basic constraints: the difficulty of applying a common model to a diverse range of developing countries, of quantifying key influences in macroeconomic performance such as structural and institutional change, of adequately capturing dynamic effects, and of identifying when the Fund's influence is exerted on macroeconomic performance.

I. Introduction

The means of evaluating Fund-supported programs are the subject of perennial controversy, at the heart of which are two broad issues. First, what sorts of questions should be asked in such evaluations? In general, reviews have focussed on one or more of three broad, but distinct, questions: are Fund-supported programs designed effectively to address countries' macroeconomic problems; are such programs adequately implemented; and do such programs have significant independent effects--that is, do they bring about developments significantly different from those that would have occurred in the absence of Fund support? A second critical issue for program evaluation is the methodology used to address the chosen questions. Clearly, the methodology needs to be tailored to the question being addressed, but the range of possibilities is large: at one end of the spectrum are individual case studies employing anything from simulation models to descriptive analyses; at the other end are aggregate exercises based on comparisons of data from before and after programs, of data from program countries and another control group, of actual data and a hypothetical counterfactual, and of target and actual policies and outcomes.

A recent staff study of programs supported through the Enhanced Structural Adjustment Facility (ESAF) has renewed interest in these issues. 1/ The study was focussed mainly on answering the first of the three questions identified above for motivating evaluations of Fund-supported programs: specifically, it examined whether the policies and outcomes during SAF/ESAF arrangements conformed to the objectives of the ESAF and why some countries were more successful than others in progressing toward the ESAF objective of external viability, satisfactory growth, and low inflation. It relied heavily on before-after comparisons of economic developments for the average of all countries, subgroups of the countries and individual countries. The study generated a lively debate (see Killick (forthcoming)) about other questions that might have been addressed. One such question, on which attention has been focussed in some other reviews of IMF- or World Bank-supported programs, is how much the presence of Fund or Bank support improved the outcomes of important macroeconomic target variables from what they would have been in the absence of such support.

The purpose of this paper is to re-examine the data from ESAF-supported programs approved during the first six years of the facility's operation with the aim of identifying the independent effects of the programs on important macroeconomic variables--growth, inflation, and a key indicator of

---

1/ The study (Schadler et al. (1993)) reviews the experience of the 19 countries that had ESAF arrangements approved before mid-1992. The ESAF, described fully in the paper, is the Fund's lending facility for supporting, at concessional interest rates, low-income countries undertaking three-year programs of macroeconomic adjustment and structural reform.

progress toward external viability (the external debt service ratio). 1/ In order to do this, it is necessary to construct a policy counterfactual (that is, policies that would have been implemented in the absence of Fund support) and develop a framework for quantitatively differentiating the effects of the counterfactual policies, exogenous developments (for example, terms of trade changes or weather), initial conditions, and the involvement of the Fund. A methodology incorporating these elements was developed by Goldstein and Montiel (1986) by adapting techniques from the literature on labor training evaluation. Applications of the technique, referred to as the General Evaluation Estimator (GEE), can be found in Greene (1989), Khan (1990), Corbo and Rojas (1992), and Faini, De Melo, Senhadjo and Stanton (1991); for the most part, these studies have adhered closely to the GEE as proposed in Goldstein and Montiel. In this study, the application of the GEE also broadly follows Goldstein and Montiel to facilitate comparisons of results with those of other studies and because data constraints thwarted most efforts to elaborate on the basic framework.

A second objective is to investigate the robustness of the estimates of the effects of Fund-supported programs using the GEE for the sample of ESAF-eligible countries and to experiment with some techniques for easing the restrictive assumptions in the GEE. The GEE, as it has been applied in virtually all earlier studies, requires that two basic conditions be met: first, that a single, relatively simple macroeconomic model reasonably captures the interaction between macroeconomic policies and outcomes for a large number of countries over time; and second, that the counterfactual to policies during periods when Fund-supported programs are in place can be characterized by a policy reaction function estimated for periods when and in countries where Fund support is not in place. The paper devotes considerable emphasis to testing these basic propositions. This effort is important because conceptually the GEE is superior to other methodologies that have been used for identifying the independent effects of Fund-supported programs. The scope for reliably quantifying the effects of Fund support, therefore, is contingent on the validity, for any sample, of the assumptions underlying the GEE.

II. Specification of the Model

1. The basic model

The GEE is geared toward answering the question "Did the involvement of the IMF through an ESAF arrangement significantly improve the outcomes for important macroeconomic variables relative to what they would have been in the absence of ESAF-support?". To answer this question the macroeconomic

---

1/ As such, the exercise reported in this paper, while complementary to that in Schadler et al., is completely separate from it. The two exercises address quite different questions and their results do not provide comparable assessments of Fund-supported programs.

outcomes or target variables in countries are described as a function of: (i) policies that would have been observed in the absence of a Fund-supported program; (ii) exogenous external factors; (iii) the existence or otherwise of a Fund-supported program; and (iv) unobservable random shocks:

$$y_{ij} = \beta_{oj} + \beta_{jk}x_{ik} + \alpha_{jh}w_{ih} + \beta_j^{IMF}d_i + \epsilon_{ij} \tag{1}$$

where y_{ij} is the jth target variable in country i, x_{ik} is a k-element vector of policy variables that would be observed in country i in the absence of Fund support, w_{ih} is an h-element vector of exogenous external variables for each country i, d_i is a dummy variable equal to one if a Fund program is in effect and zero otherwise, and ϵ_{ij} is a zero mean, fixed variance, serially uncorrelated error. For the jth target variable, β_{jk} and α_{jh} are kxl and hxl vectors, respectively, of fixed parameters. After postulating a rule for policies in the absence of a Fund-supported program (x_{ik}), the model is estimated using pooled cross-section and time-series data drawn from countries and periods in which Fund support was in place and those in which Fund support was absent. The aim is to get consistent estimates for β_j^{IMF}, the "independent effect" of Fund-supported programs on each target variable. If these are statistically significant at a reasonable confidence level, Fund-supported programs are found to have significant effects.

Policies adopted in the absence of a Fund-supported program (x_{ik}) are directly observable only for nonprogram periods, and thus a key element of the GEE is the construction of a counterfactual for policies during programs. In Goldstein and Montiel (1986) and subsequent empirical applications, this counterfactual is based upon a policy reaction function that links changes in policy instruments to the deviation of the observed lagged value for each target from its desired value, y_{ij}^d. Specifically, the policy reaction function is described by:

$$\Delta x_{ik} = \gamma_{kj}\left[y_{ij}^d - y_{ij(-1)}\right] + \eta_{ik} \tag{2}$$

where y_{ij} is a j-element vector of target variables, η_{ik} is a zero mean, fixed variance, serially uncorrelated error term assumed to be uncorrelated with ϵ_{ij}, and Δ is the first difference operator. 1/ The kxj parameter matrix γ_{kj} indicates the extent to which policy instruments are adjusted in response to disequilibria in the target variables.

Substituting equation (2) into equation (1) and subsuming y_{ij}^d in the constant (that is, assuming y_{ij}^d is invariant across countries and over time) this substitution gives: 2/

1/ Implicitly, the lack of mutual correlation between η_{ik} and ϵ_{ij} implies that changes in policy instruments (Δx_{ik}) are not influenced by contemporaneous exogenous shocks to the target variables.

2/ By assumption the stochastic structure is one in which shocks are transitory, thereby ruling out a source of bias in equation (3) because of nonzero correlation between the error terms and explanatory variables.

$$\Delta y_{ij} = \beta_j^o - (\beta_{jk}\gamma_{kj} + 1)y_{ij(-1)} + \beta_{jk}x_{ik(-1)} \tag{3}$$
$$+ \alpha_{jh}w_{ih} + \beta_j^{IMF}d_i + (\epsilon_{ij} + \beta_{jk}\eta_{ik})$$

Equation (3) constitutes the basic GEE reduced form model as applied in most earlier studies. While conceptually rigorous, its operational usefulness depends on the validity of several rather restrictive assumptions. The remainder of this section considers some of these issues and suggests possible ways of addressing them. 1/

2. The counterfactual--which policy reaction function?

The choice of a policy reaction function is a critical step in applying the GEE, and the theoretical and empirical literature offers a profusion of possibilities. The policy reaction function can be set explicitly in an optimizing framework, but this does little, if anything, to narrow the choice; different specifications of policymakers' objective functions and of the constraints they face generate a wide range of alternative reaction functions (see Turnovsky (1977)). This paper starts with the specification of the policy reaction function used by Goldstein and Montiel (1986), Greene (1989) and Khan (1990) (equation (2)) and experiments with two alternative less restrictive specifications, which are described in this sub-section.

According to equation (2), policies are adjusted only in response to target disequilibria in the previous period. However, it is plausible that policy adjustments in the current period also take account of the impact of external exogenous factors not captured in the lagged value of the target variable. Changes in the terms of trade or in primary commodity prices may have lagged effects; for example, fiscal revenues this year may be weakened by a drop in export prices last year. Assuming policymakers adjust policies with information only on past exogenous shocks (that is, policies are set at the beginning of the year), an enhanced policy reaction function with exogenous external factors (w_{ih}) lagged one period gives:

$$\Delta x_{ik} = \gamma_{kj}\left[y_{ij}^d - y_{ij(-1)}\right] + \mu_{ih}w_{ih(-1)} + \eta_{ik} \tag{4}$$

When equation (4) is substituted into equation (1) to eliminate the unobservable values of x_{ik} (and subsuming y_{ij}^d in the constant) the specification of the GEE becomes:

$$\Delta y_{ij} = \beta_j^o - (\beta_{jk}\gamma_{kj} + 1)\, y_{ij(-1)} + \beta_{jk}\, x_{ik(-1)} + \alpha_{jh}w_{ih} + \beta_{jk}\mu_{ih}w_{ih(-1)}$$
$$+ \beta_j^{IMF}d_i + (\epsilon_{ij} + \beta_{jk}\eta_{ik}) \tag{5}$$

A second variation of equation (2) results from setting the formulation of policies explicitly in an optimizing framework. One simple approach is

1/ Some of these restrictive assumptions characterize other, especially cross-section, estimators.

to assume that policymakers in each country i choose policies (x_{ik}) to minimize a quadratic loss function (L) of the form

$$L = (y_{ij} - y^d_{ij})^2 \tag{6}$$

Subjecting this to the constraint of the economic model postulated in the GEE framework,

$$y_{ij} = \beta_{oj} + \beta_{jk}x_{ik} + \alpha_{jh}w_{ih} + \epsilon_{ij} \tag{7}$$

the policy reaction function takes the form in levels

$$x_{ik} = \hat{\gamma}_o + \hat{\gamma}_{kh}w_{ih} \quad \underline{1}/ \tag{8}$$

with y^d_{ij} (assumed to be fixed) and other constant terms subsumed in $\hat{\gamma}_o$. After substitution for x_{ik} in equation (1), the reduced form GEE becomes

$$\Delta y_{ij} = \hat{\beta}^o_j - y_{ij(-1)} + (\beta_{jk}\hat{\gamma}_{kh} + \alpha_{jh})\, w_{ih} + \beta^{IMF}_j d_i + \epsilon_{ij} \tag{9}$$

In this equation, terms in x_{ik} drop out and the interpretation of the estimated coefficients on the lagged target variable and the exogenous external factors differs from that in equation (3). 2/

In practice, in an exercise where the goal is to evaluate the effectiveness of Fund-supported programs, an appeal to optimizing principles to motivate the policy reaction function may be thought to create a potential conundrum for the counterfactual approach. If countries are assumed to adopt optimal policies in the absence of a Fund-supported program, why would a country ever turn to the Fund for support? Following this line of argument, the influence of a program would depend on the Fund's seal of approval strengthening confidence in the economy (and catalyzing external assistance) and thereby enhancing the effectiveness of any given stance of policy. Yet, if policies are the same with or without a program, this presumes that domestic and foreign economic agents fail to recognize the equivalence of policies under the two regimes. In reality the Fund's

---

1/ This form of the policy reaction function is derived by assuming that $E(\epsilon_{ij}) = 0$, and that the covariances between the parameters β_{jk} and ϵ_{ij}, and between α_{jh} and ϵ_{ij}, are equal to zero in the analytical derivation.

2/ To derive equation (2) from an optimizing framework, equation (6) must be maximized subject not to the economic model postulated in the GEE (i.e., equation (7)), but rather to the constraint

$$\Delta y_{ij} = b_{jk}\Delta x_{ik} + u_{ij}$$

where the parameters b_{jk} are assumed to be stochastic and u_{ij} is a random disturbance. Also necessary are the simplifying assumptions that $E(u_{ij}) = 0$ and that the covariance between b_{jk} and u_{ij} is set equal to zero in the analytical derivation.

seal of approval expands the opportunity set of policies beyond that available to a country not receiving Fund support. In fact, almost all of the countries that had ESAF arrangements were also seeking agreements with creditors to reschedule or restructure external debt, for which entering an arrangement from the Fund was a prerequisite.

Even with careful specification of the policy reaction function, there remains a question of whether individual country behavior can be sensibly aggregated in a uniform model that is stable across countries and over time. Specifically, differing institutional characteristics (for example, the degree of policy discipline inherent in specific exchange rate arrangements or the relationship with a major donor), changing political conditions, or varying severities of economic distress are likely to result in countries formulating policies with respect to different or changing objective functions or subject to different or changing constraints. Another question is whether it is appropriate to assume that the policy reaction function of a program country had it not received Fund support is identical to that of a nonprogram country that did not seek Fund support. For example, the counterfactual for a country receiving Fund support may involve the imposition of exchange restrictions, while countries that do not seek Fund support may constrain themselves to "Fund-type" policies--that is, avoiding the use of exchange restrictions.

3. Characterizing the independent effect of Fund-Supported programs

The simple GEE framework captures the independent effect of Fund support in an additive term ($\beta_j^{IMF} d_i$) constant across countries and over time. This term is meant to capture four channels through which a Fund-supported program could affect the macroeconomic targets: (i) changes in the state of confidence in the economy; (ii) changes in the desired value of targets, for example through structural reforms aimed at raising the rate of potential growth; (iii) policies different from what they would have been in the absence of a program; and (iv) changes in the effectiveness of any given stance of policies.

Ideally, the specification of the effect of Fund support would not be restricted to an additive term constant across countries and time but would decompose program effects in a way allowing their magnitude and type to vary across programs. In doing so, it is reasonable to assume that the first two channels described above--changes in confidence and in desired values of targeted variables--could be captured in additive terms. Even here, however, it would be optimal to allow the coefficient β_i^{IMF} to vary across program years (that is by allowing program-specific coefficients).

To capture the third channel of program influences--policies different from what they would have been in the absence of Fund-support--modifications to the model are needed. 1/ The policies supported by the Fund differ for

1/ Much of this discussion draws upon Goldstein and Montiel (1986).

each country, reflecting countries' preferences on how to address their problems, the nature of the macroeconomic disequilibria, and the severity of the external financing constraint. Thus, for given relationships between policies and targets, the effect of Fund-supported programs on macroeconomic targets should differ across countries and over time. This variation could be captured with multiplicative dummies to relate the estimate of program effects directly to the difference between policies under the program and the counterfactual. The effect of a Fund-supported program would be

$$\beta_{ijk}^{IMF} = \beta_{jk}\left(x_{ik}^{P} - x_{ik}\right) + A_{j}^{IMF} \tag{11}$$

where x_{ik}^{P} is the vector of (observed) policies adopted under a program, and A_j^{IMF} is a constant. The first term in (11) captures the effect of changes in policy instruments and A_j^{IMF} all other program effects.

The fourth type of program effect--changes in policy effectiveness--operates by altering the β_{jk} parameters, for given changes in policy variables. For example, changes in fiscal and credit policy under a Fund-supported program may have a greater impact than they would otherwise have had owing to credibility effects when policies and targets are set in a pre-announced medium-term framework. In these circumstances, estimates of equation (3) that assume constancy of the β_{jk} coefficients are open to the "Lucas critique" (Lucas (1976)), and may not measure the true value of β_j^{IMF}. Ideally, the variation of the β_{jk} parameters would be endogenized by modelling the link between public policies and private expectations. Alternatively, the β_{jk} coefficients could be allowed to vary across countries and over time, or the variation in the β_{jk} coefficients could be restricted to systematic differences between program and nonprogram periods. Under the latter specification, the effect of Fund support is:

$$\beta_{ij}^{IMF} = \beta_{jk}^{P}\, x_{ik}^{P} - \beta_{jk}x_{ik} + A_{j}^{IMF} \tag{12}$$

or

$$\beta_{ij}^{IMF} = \beta_{jk}\left(x_{ik}^{P} - x_{ik}\right) + \left(\beta_{jk}^{P} - \beta_{jk}\right) x_{ik}^{P} + A_{j}^{IMF} \tag{12a}$$

of which expression (11) is a special case and β_{jk}^{P} is a vector of parameters linking policies and targets during programs. The second term in equation (12a) captures the effect of changes in policy effectiveness.

Two steps are needed to get estimates of β_j^{IMF} as defined in equation (12) or (12a). First, the policy reaction function must be estimated to get values of x_{ik} for program periods (the unobservable counterfactual policies). Second, using the estimates of x_{ik} and substituting for β_j^{IMF} in equation (1), the following equation must be estimated.

$$y_{ij} = \beta_{oj} + \beta_{jk}x_{ik}(1-d_i) + \beta_{jk}^{P}\, x_{ik}^{P}d_i + \alpha_{jh}w_{ih} + A_j^{IMF}d_i + \epsilon_{ij} \tag{13}$$

In sum, plausible characterizations of the influence of programs on targets suggest that a simple invariant additive term in the GEE may not do full justice to the range of potential program effects. In practice, the possibility of a more informative decomposition of program effects that allows variation across countries requires a large sample and the ability to identify empirically a stable policy reaction function as examined in Section IV.

4. Dynamics and the importance of initial conditions

Dynamic characteristics of the economy are likely to influence the effects of Fund-supported programs in two main ways that are captured to varying degrees in the simple GEE model. First, the full effects of changes in policies--those supported by the Fund or the counterfactual--and exogenous influences on target variables, particularly output growth, may occur with long lags. In the simplest form of the GEE presented in this paper, *such dynamic effects are not allowed for: effects from changes in* policies and exogenous factors are assumed to occur within one year. Some applications (Khan (1990)) have eased this restriction by estimating the model with two year average (rather than single year) data for the target variables. This procedure goes some way toward accounting for short lags, but at the cost of not accounting for contemporaneous effects of policies in the second year, which may or may not be a program year.

A second role for dynamic influences works through initial conditions in two ways. First, departures from the desired values of the macroeconomic targets in the preprogram period may generate a policy response even in the absence of a program. In the GEE model, this influence is captured by the term $\beta_{jk}\gamma_{kj}$ in the coefficients on the lagged target variables in equation (3). If the likelihood of adopting a program is positively correlated with negative prior shocks, then a failure to account for corrective policy responses that would have taken place in the absence of Fund support would lead to an overstatement of the true effect of a Fund-supported program.

Second, current values of the target variables may be influenced by initial conditions through inertial effects. 1/ These are not captured in the static model postulated in equation (1), where the stochastic terms ϵ_{ij} and η_{ik} are assumed to be serially uncorrelated: all shocks are assumed to be transitory and to cause one-period changes in target variables that are fully reversed in the following period. This is a quite restrictive assumption. Not only does it require that the time series for each target variable be stationary, but it also rules out a wide range of stationary stochastic processes for which the impact of temporary shocks persists over

---

1/ Inertial effects may arise for a variety of reasons, such as backward-looking indexation, slowly-adjusting expectations, staggered contracts, and transaction costs.

time. 1/ If, in fact, significant inertia exists, then imposing full one-period reversion to mean will result in an understatement of the positive effects of a Fund-supported program after a negative shock.

In principle, the following more general dynamic form of equation (1) would allow a wider range of potential dynamic effects:

$$y_{ij} = \rho_s(L)^{-1}\beta_{oj} + \rho_s(L)^{-1}\beta_{jk}(L)x_{ik} + \rho_s(L)^{-1}\alpha_{jh}(L)w_{ih} + \rho_s(L)^{-1}\beta_j^{IMF}d_i + \rho_s(L)^{-1}\epsilon_{ij} \quad (14)$$

where L is the lag operator and $\rho_s(L)$ is a lag polynomial in L of order s which captures inertia in the behavior of the target variables. The order of the lag operators may differ for each variable. Substituting equation (2) (the policy reaction function) into equation (14) gives:

$$\Delta y_{ij} = \beta_j^o - \left(\rho_s(L)^{-1}\beta_{jk}(L)\gamma_{kj} + 1\right)y_{ij(-1)} + \rho_s(L)^{-1}\beta_{jk}(L)x_{ik(-1)} + \rho_s(L)^{-1}\alpha_{jh}(L)w_{ih} + \rho_s(L)^{-1}\beta_j^{IMF}d_i + \left(\rho_s(L)^{-1}\epsilon_{ij} + \rho_s(L)^{-1}\beta_{jk}(L)\eta_{ik}\right) \quad (15)$$

which is a general dynamic form of the reduced form GEE (equation (3)).

III. Estimation Procedures

1. The sample

The model, as specified in equation (3) with some of the modifications suggested in Section II, is estimated for the period 1986-91 with data for 61 of the 66 countries (Appendix Table 1) eligible to use ESAF resources as of 1992 (exclusions are noted below). Nineteen of these countries had ESAF arrangements at some time during the sample period. 2/ The sample was restricted to ESAF-eligible countries, rather than a larger set of developing countries, for two reasons: first, the many years when stand-by and extended arrangements were in effect in non-ESAF-eligible countries could not be characterized as nonprogram years, yet these arrangements were

---

1/ In the absence of stationarity the concept of reversion to mean is not well defined because the mean of the stochastic process is not time-invariant, and the series will tend to move continuously away from a given level as a result of the endless impact of past and current shocks. See Harvey (1981) and Priestley (1981).

2/ For these countries, program years are those when either a SAF or ESAF arrangement was in place. SAF arrangements typically have less stringent conditionality than ESAF arrangements. For most countries that had ESAF arrangements, however, prior SAF arrangements were used to establish commitment to adjustment and were close in nature to ESAF programs.

in many respects not comparable to ESAF arrangements; second, including only low-income countries reduced the scope for parameter instability owing to differences in structural and institutional conditions between low-income countries and other developing countries.

Even the sample restricted to ESAF-eligible countries is quite diverse. The program countries are dominated by heavily indebted African countries with a narrow range of exports and relatively simple market structures. The nonprogram countries are more diverse including, in about equal proportions, indebted African countries and small Caribbean or Pacific island economies with close institutional and financial links to particular industrial countries. There are also a few South and Southeast Asian countries.

For estimation, a number of data points are dropped from the full sample of ESAF-eligible countries. Some observations are excluded because of inadequate data owing to civil strife (Afghanistan (1990-91), Angola, Liberia, and Nicaragua (all years)); major discontinuities which could not be corrected (Djibouti (all years)); extreme isolation (Albania (all years)); or political discontinuities (Democratic Republic of Yemen (1990-1991)) and Yemen Arab Republic (1989-1991)). Years in which countries had a SAF arrangement (except when followed immediately by an ESAF arrangement), Stand-by, or Extended arrangement are omitted because they were considered invalid as nonprogram counterfactuals; this, together with lack of data in 1991, exclude Somalia entirely from the sample.

The exclusion of data points renders the structure of the panel data of program and nonprogram years incomplete. Techniques for analyzing panel data in which missing observations are random or follow a regular (or "rotating") pattern are not applicable because the gaps in the data do not conform to either of these patterns. Instead, for estimation purposes, the panel data are handled as a pooling of annual observations, with the number of program and nonprogram years varying from country to country.

For the purpose of investigating the stability and the dynamic specification of the GEE a larger set of observations for each country would have been preferable. Extending the sample, however, would have required dropping several countries owing to the lack of consistent data for earlier periods. Also, the number of additional useable observations (that is, years in which countries did not have a Stand-by or Extended arrangement in place) would have been a rather small proportion of the total.

Even for the limited sample covered, the quality of data is poor. In many instances, the accuracy of the measures of macroeconomic variables, such as GDP, is likely to very weak. Also, ad hoc correction for breaks in the series were frequently needed. These fundamental weaknesses in the data qualify the inferences or judgements that can be drawn from them.

2. Definitions: targets, policies, exogenous influences and period of Fund support

This study considers three target variables (y_{ij}) that reflect the objectives of the ESAF: (i) the growth rate of real GDP; (ii) the average rate of consumer price inflation; and (iii) a measure of external viability--the ratio of external debt service to exports. The last target is preferred to other external indicators, such as the current account, the overall balance of payments, or the level of international reserves, as a measure of external viability for several reasons. 1/ Most countries entered ESAF arrangements with large debt overhangs, and reducing the debt service ratio to manageable levels was the primary external objective. Within this goal, targeted and actual outcomes for other external variables varied widely depending on initial conditions and prospects for attracting concessional inflows. Moreover, for non-program periods, developments in these other variables sometimes reflected the imposition of formal or informal trade and exchange restrictions rather than changes in the viability of the external position. Increases in reserves were at times associated with the accumulation of arrears. Also, a reserves target did not exist for a sizable proportion of the countries (CFA and ECCB country members) that did not directly own international reserves

Three policy instruments (x_{ik}) are considered: (i) the deficit of the central government in relation to GDP; (ii) the growth of net domestic assets of the banking system (NDA) 2/; and (iii) the change in the nominal effective exchange rate (NEER). Ideally, the vector of policy instruments would also include indicators of structural reforms/conditions and institutional arrangements (such as flexible or fixed exchange rate regimes). However, these variables cannot be easily quantified or reduced to an index. The external environment indicators (w_{ih}) comprise changes in the terms of trade and the growth of export markets. 3/ Whenever possible, the data were taken from Executive Board documents.

Several difficulties arise in defining the variable denoting the presence of a Fund-supported program (d_i). First, the distinction between program and non-program years is blurred when Fund support starts in the middle of the year. For this study, any year in which a SAF/ESAF program

---

1/ The debt-service ratio is not, however, an infallible indicator of progress toward external viability. For example, changes in this ratio may at times only reflect a temporary variation in export prices.

2/ In principle, the controllable monetary policy instrument of the authorities is net domestic assets (or credit) of the central bank. However, data on a comparable basis across countries were not available.

3/ In principle, only one of market price and volume indicators should be used. In practice, the suitability of each varies among the countries: the terms of trade are relevant for small primary producers, but world market growth is relevant for large primary producers and countries with differentiated exports.

was in effect for six months or more was considered a program year. In applying this rule, the program periods (periods over which a program was framed) rather than annual arrangement periods (periods over which programs were approved) were considered. Even this rule, however, does not clearly delineate the period during which Fund support influenced policies and outcomes. Usually, substantive negotiations and policy actions occurred in anticipation of Fund support in the year preceding the formal program. Also, in some cases, the Fund had a considerable influence on policies after an ESAF arrangement. For example, The Gambia's SAF/ESAF arrangements stretched from FY1986 (July-June) to FY1990, but even in FY1991, the Fund monitored macroeconomic developments vis-a-vis quantified targets agreed with the authorities.

How to handle variations in program implementation across countries presents a second difficulty. Some would argue that implementation should be reflected in the estimated model, for example, by representing the influence of the Fund as the proportion of purchases made relative to total access under the arrangement. Purchases are likely to be an imperfect indicator of implementation, however, because purchases in SAF/ESAF programs are scheduled at low (six-month) frequencies and waivers may be granted to permit purchases even in the event of policy slippages. Alternatively, the effectiveness of Fund support could be judged in terms of outcomes, whether or not programmed policies are implemented: a period when a country has an agreed program with the Fund but fails to implement it or meet the targets is treated as a program year when the effect of the Fund is zero. This study takes the latter approach and uses a binary, one-zero index of Fund involvement for the dummy variable (d_i).

A third point to recognize is that estimates of Fund-supported programs will likely include the effects of parallel World Bank programs. As distinct from the Fund's Stand-by and Extended Facilities, SAF/ESAF arrangements require explicit collaboration among country authorities, the World Bank and the Fund.

IV. Results

1. The generalized evaluation estimator

Estimation results were obtained for the basic GEE equation (3) and several of the modified versions suggested in Section II to ease the restrictiveness of the simple framework: (i) the expanded equation (5), which includes lagged exogenous external influences in the policy reaction function was estimated; 1/ (ii) country and time dummies were introduced to help account for some of the cross-country differences in economic

---

1/ The policy reaction function resulting from the optimization subject to the constraint implied by equation (1) was not tried because it represented a substantial simplification of even equation (3).

structures and time-specific exogenous developments not captured in the terms of trade and market growth variables; and (iii) a correction procedure developed by Heckman (1979) (explained in detail in the Annex) was tried to correct for possible sample selection bias. Other methods discussed in Section II to ease the restrictive assumptions of the GEE were not feasible. Specifically, consideration of a richer characterization of the effects of Fund support as described in equations (11-13) was not possible because the estimates of the policy reaction function, which are needed to estimate counterfactual policies during program periods, proved to be very poor. Also, because of the short time series available, a more complex dynamic structure for the GEE could not be explored.

Regression estimates based on pooled time-series, cross-country data are prone to heteroschedasticity, and even after the inclusion of country-specific and time dummies in the estimated equations, regression residuals displayed heteroschedasticity. 1/ A weighted least squares correction for heteroschedasticity was not attempted because, without information on the form of heteroschedasticity, the primacy of one weighing scheme over another is unknown. 2/ Instead, the ordinary least squares coefficients were retained, and the reported t-statistics were computed from heteroschedastic-consistent estimates of the standard errors based on White's variance-covariance estimator that provides consistent estimates even when the exact form of heteroschedasticity is not known. 3/

Table 1 presents the preferred estimates from these exercises. These estimates exclude the lagged values of the exogenous variables (the modification suggested in equation (5)) because their coefficients were not significantly different from zero and the terms had little effect on the fit of the regression or the estimates of the other coefficients. The results also exclude the time dummies mentioned above because their coefficients were generally not significant and had little effect on (or worsened) the fit of the equations. The impact of Fund-supported programs is found to be sizeable and statistically significant with respect to growth (at the 5 percent level) and the external debt service ratio (at the 10 percent level), but not inflation. On average, growth rates are found to be more than 1 percentage point per annum higher during program periods than they would have been in the absence of a Fund-supported program. Debt-service

---

1/ Statistically significant values of the Breusch-Pagan test for heteroschedasticity (Breusch and Pagan (1980)) were observed in the estimated GEE equation for inflation (at the 5 percent significance level) and the external debt service ratio (at the 1 percent level).

2/ With only one nonprogram observation for several countries, there are insufficient degrees of freedom to use the common weighted least squares procedure in which observations for each country are weighted by the inverse of the standard deviation of the corresponding estimated residuals.

3/ See White (1990). The properties of White's heteroschedastic-consistent variance-covariance estimator are conditional upon a correct specification of the estimated equation up to an additive error term.

Table 1. Estimates of the GEE 1/

Target Variable	Real GDP Growth Rate	Inflation Rate	External Debt Service Ratio
Constant	-6.619 (-1.71)	10.248 (1.08)	22.258** (3.98)
Lagged real GDP growth rate	-1.107** (-17.96)	-0.764* (-2.18)	0.022 (0.09)
Lagged inflation rate	0.0005 (0.13)	-0.687** (-4.76)	0.027 (1.09)
Lagged external debt service ratio	0.013 (0.74)	0.106 (1.14)	-0.376** (-3.09)
Lagged fiscal balance/GDP	-0.042 (-1.37)	-0.467 (-1.31)	0.097 (0.76)
Lagged net domestic asset growth	0.004 (1.82)	-0.088 (-1.47)	-0.020 (-1.78)
Lagged percentage change in NEER	-0.009 (-1.03)	0.436* (2.12)	0.058 (1.05)
Current percentage change in terms of trade	0.002 (0.21)	-0.104 (-0.78)	-0.104** (3.44)
Current export market growth	0.090 (1.78)	0.293 (1.26)	-0.059 (-0.30)
IMF program dummy	1.374* (2.18)	-3.330 (-0.35)	-5.552 (-1.75)
$\bar{R}^2$	0.537	0.398	0.069
S.E.E.	3.259	29.612	15.734
Number of observations	291	291	291
Breusch-Pagan test for heteroschedasticity	1.35	10.83*	23.71**
Jarque-Bera test for normality of residuals	26.57**	28,231.00**	7,086.90**

1/ The regression estimates were obtained using a ordinary least squares procedure, with country-specific dummies included in the specification. Standard errors and t-statistics of coefficients are computed using White's heteroschedasticity-consistent variance-covariance estimator. The figures in parentheses are t-statistics; $\bar{R}^2$ is the adjusted coefficient of determination; S.E.E. is the standard error of the regression. A single asterisk indicates statistical significance at the 5 percent level; two asterisks indicate statistical significance at the 1 percent level.

ratios are found on average to be more than 5 percentage points lower during program periods than they would have been in the absence of a Fund-supported program. This effect on the debt-service ratio is not attributable to the link between Paris and London Club agreements and Fund support, because the debt service ratio is measured before debt relief. 1/

On the other right hand-side variables very few coefficients are significant at the 5 or 1 percent confidence level. The exceptions are own lagged levels of the target variables, which enter with coefficients significantly different from zero at the 1 percent level; the lagged level of the real GDP growth rate and the change in the nominal effective exchange rate (which has an implausible positive sign) in the inflation equation; and the terms of trade in the debt service ratio equation. The lagged fiscal balance and lagged net domestic asset growth are not found to have a significant impact on the outcome of any target variable. 2/

In broad terms, these results are similar to those of Khan (1990)--the most recent and largest study using the GEE methodology to estimate the effects of programs supported by upper credit tranche Stand-by and Extended arrangements. That study uses a sample of 1,104 observations spanning 69 countries over the period 1973-88, including 315 program years. 3/ It is noteworthy that Khan's sample is quite different from that used in this study: it excludes SAF/ESAF arrangements from the program sample, ends several years prior to the sample used in this paper and includes low- and middle-income developing countries. Like this study, Khan finds that lagged values of target variables significantly influence current changes of these variables. He also finds few significant effects of lagged policies on current target variables; the main difference in this respect is that Khan finds significant and plausible effects of the lagged fiscal balance on the target variables, while in this study the fiscal effect is found to be insignificant. Both studies find programs lead to reductions (though not statistically significant) in inflation. The most important difference

---

1/ The effect of stock of debt reduction operations associated with Fund-supported programs would be reflected in measures of the debt service ratio before debt relief in years subsequent to the debt reduction. However, in the sample of program countries considered in this study no stock of debt operations linked to Fund-supported programs were undertaken.

2/ The residuals of most of the estimated equations fail to pass the Jarque-Bera test for normality. T-tests should, therefore, be interpreted cautiously as they may be sensitive to nonnormality in a fashion that is determined by the numerical value of the regressors. This cautionary note applies also to other regression diagnostic tests (see Jarque-Bera (1987)).

3/ Khan estimates the GEE for a slightly different array of economic variables. Targets comprise the real GDP growth rate, the inflation rate, the ratio of the current account to GDP, and the ratio of the balance of payments to GDP. Policy variables are the growth of domestic credit, the ratio of the fiscal balance to GDP, and the real effective exchange rate. A time trend and the change in the terms of trade were also included.

between the two studies, however, is in the estimate of the effect of Fund support on growth: Khan finds a significantly negative effect on growth, in contrast to the significantly positive effect in the current study.

2. The policy reaction function

Identification of the coefficients and measurement of their standard errors in the policy reaction function from the reduced form GEE is neither possible nor necessary for obtaining estimates of the β_j^{IMF} terms. 1/ Nevertheless, obtaining estimates of them, by directly estimating the counterfactual policy reaction functions with data from the (observable) nonprogram periods only, is useful for two reasons. First, it provides a means of evaluating the validity of the reaction functions for the sample over which they can be estimated. Second, direct estimates of the policy reaction function are part of the procedure to correct for sample selection bias arising when unobservable factors that influence countries' decisions to receive Fund support also influence their policy reactions. To identify whether this potential source of bias exists, and to correct for it, a two-step procedure first proposed by Heckman (1979) is used. 2/ First, a probit model of the probability of not having a program is estimated. Second, the probit estimates are used to calculate the inverse Mills ratio 3/ (for each nonprogram observation) which, when included in the policy reaction function estimated for the nonprogram observations, accounts for the bias due to nonrandom sampling. Annex I describes the procedure in detail and reports estimates of the probit model.

1/ The γ_{kj} parameters cannot be identified from the single equation estimates because the number of structural parameters exceeds the reduced form coefficients. However, by pooling the parameter estimates from the three equations the γ_{kj} parameters can be identified if policy instruments and macroeconomic targets are equal in number. When the number of instruments is less than that of targets the γ_{kj} parameters cannot be identified, and when they exceed the number of targets, multiple solutions exist. It is not possible to measure the standard errors of the estimates. For the alternative specifications of the GEE (equations (5) and (15)), which introduce lagged exogenous variables and an unrestricted serial correlation coefficient on lagged target variables, the γ_{kj} parameters are not identifiable regardless of the number of targets and instruments.

2/ *Goldstein and Montiel (1986), Greene (1989), and Khan (1990) do not* attempt to correct for this potential source of bias, but several studies of World Bank adjustment lending do (see World Bank (1990), Faini et al. (1991), and Corbo and Rojas (1992)).

3/ The inverse Mills ratio is $\phi(-\delta' D_i)/[1-\Phi(-\delta' D_i)]$ where ϕ and Φ are, respectively, the density and distribution function for a standard normal variable; D and δ are a vector of explanatory variables and parameters, respectively, in the probit model. The ratio is a monotone decreasing function (ranging from 0 to ∞) of the probability that an observation is selected into the sample of nonprogram countries, $[1-\Phi(-\delta' D_i)]$.

As with the GEE, country and time dummies were introduced in the policy reaction functions. Most of the time and country dummies had coefficients insignificantly different from zero and had little effect on or worsened the overall fit of the equation (reduced the $\bar{R}^2$) and therefore were not retained in the reported results.

The regression estimates of the policy reaction function are poor in several respects (Table 2): the $\bar{R}$-squared statistics are negative or very close to zero; t-statistics for individual coefficients are insignificant (except on the debt service ratio in the equation for the nominal effective exchange rate); F-tests can not reject the null joint hypothesis of zero slopes; and the regression residuals exhibit statistically significant heteroschedasticity and nonnormality. The insignificant coefficients of the inverse Mills ratio suggest that sample selection bias is not present in the sample. However, in view of the poor performance of the estimated policy reaction function this result is not particularly revealing. In short, these estimates provide a weak basis for deriving estimates of the unobservable counterfactual policies for program periods. Thus, no attempt was made to estimate equation (13) to identify more complex characterizations of the independent effects of Fund-supported programs.

3. Significance and stability of the estimates

How much confidence can be placed on the estimates of the effects of programs? The applications of the GEE to date have tended to focus on the size and significance of the estimates of the β_j^{IMF} parameters, without testing the validity of the basic assumptions in the GEE or, in some cases, even without reporting indicators of the overall goodness of fit. In this section, some simple experiments to determine the robustness of the parameters estimates are reported.

There are many ways to evaluate the regression estimates, and the approach taken here is not intended to be exhaustive. As measured by the $\bar{R}$-squared statistic, the overall fit of the estimated equations is modest (almost nil for the external debt service ratio). This, together with the evidence of heteroschedastic residuals (even after the inclusion of country and time dummies) and the large number of coefficients insignificantly different from zero or with counterintuitive signs, suggests the possibility of misspecification and, therefore, of biased coefficient estimates. One potentially important omission is structural conditions/reforms that figured prominently in SAF/ESAF-supported programs. A second concern is the possibility of heterogeneity bias (Hsiao (1986)). 1/ This arises when parameter estimates are obtained by imposing identical coefficients on

1/ The key problem in heterogeneity bias is that the imposition of identical parameters leads to an averaging of coefficients that differ greatly across countries (or time) and therefore produces nonsensical results. In effect, the assumption of a "representative agent" that can be described by an average is not valid.

Table 2. Estimates of the Policy Reaction Function 1/

Policy Variable	Δ Fiscal Balance/GDP	Δ Net Domestic Asset Growth	Δ Percentage Change in NEER
Constant	1.643 (0.67)	-2.209 (-0.11)	-2.607 (-0.35)
Lagged real GDP growth rate	0.024 (0.19)	-1.090 (-1.11)	-0.204 (-0.69)
Lagged inflation rate	0.006 (1.02)	-0.081 (-0.12)	0.017 (0.29)
Lagged external debt service ratio	-0.0007 (-0.04)	-0.097 (-0.40)	-0.152* (-2.22)
Inverse Mills ratio 2/	-3.911 (-0.65)	16.271 (0.24)	13.070 (0.76)
$\bar{R}^2$	-0.013	-0.016	0.019
S.E.E.	7.064	106.339	23.239
Number of observations	203	203	203
F-statistic (zero slopes)	0.36	0.22	1.96
Breusch-Pagan test for heteroschedasticity	0.70	62.27**	21,88**
Jarque-Bera test for normality of residuals	4,875.06**	17,986.30**	495.36**

1/ The regression estimates were obtained from the sample of nonprogram observations using an ordinary least squares procedure. Standard errors and t-statistics of coefficients are computed using White's heteroschedasticity-consistent variance-covariance estimator. The figures in parentheses are t-statistics; $\bar{R}^2$ is the adjusted coefficient of determination; S.E.E. is the standard error of the regression. A single asterisk indicates statistical significance of the 5 percent level; two asterisks indicate statistical significance at the 1 percent level.

2/ Values of the inverse Mills ratio were computed using the estimated probit equation reported in Annex Table 1.

pooled data but the true values of the parameters differ significantly across countries and/or over time, owing for example to differences in policy regimes. In these circumstances, pooled estimates that ignore parameter heterogeneities are likely to be biassed. A third concern is that the regression residuals fail to pass the Jarque-Bera test for normality, signalling the risk of invalid inferential statements, even in large samples (Jarque and Bera (1987)).

The reliability of the parameter estimates is also revealed by the stability properties of the model. Earlier applications of the GEE have not reported tests of stability. 1/ Most investigate changes in program effectiveness as measured by the β_j^{IMF} coefficients (or the equivalent for World Bank programs) between two sub-periods of the sample used; Greene (1989) and Corbo and Rojas (1992) also report changes in estimates of the other coefficients of the reduced form GEE. Yet, in each study, the evidence of changes in point estimates of the β_j^{IMF} and other coefficients (at times statistically significant) is not seen as a possible indication of instability in the underlying model.

The most appealing approach to exploring instability would be to estimate a varying parameters model in which estimated parameters are allowed to vary across countries 2/ and over time 3/ and therefore also between program and nonprogram periods. This general approach has the merit of permitting tests of stability and uniformity restrictions nested within a less restrictive framework. 4/ However, such an approach requires a data set considerably larger than that available in this study: to explore inter-country instability, the number of data points for each country must exceed the number of regressors; and to explore instability between program and nonprogram periods, the number of data points for each country in each regime (program and nonprogram) must exceed the number of regressors.

In light of the constraints imposed by the structure of our sample (that is, the small and unequal number of annual observations for each country), the stability of the individual parameters β_{jk} and β_j^{IMF} is

---

1/ To the extent that the high estimated standard errors found in this study are a general feature of GEE estimates, measuring parameter instability may not be very productive because the data fit the model too imprecisely for stability tests to have much power.

2/ See Swamy (1970).

3/ See Hsiao (1986).

4/ For example, the presence of heteroschedastic errors can be taken into account in estimation and testing of coefficient variation; this property is particularly appealing in pooled cross-section, time-series data.

examined by recursive regression methods. 1/ Two types of recursive exercises were done: in the first, the reduced form GEE was estimated recursively, starting with a baseline sample (comprising all nonprogram observations plus one program observation) and then adding in program observations one at a time (reported as "program recursions"); 2/ in the second, the recursive procedure was carried out by starting with the full sample and then subtracting nonprogram observations one-by-one (reported as "nonprogram recursions"). The share of estimates in the recursions that differs significantly from the pooled sample estimates provides an indication of the sensitivity of the estimated parameters to variations in the sample. Also of interest, particularly for parameters where whole sample estimates are significantly different from zero, is the share of recursion estimates that differ significantly from zero. 3/

The recursion estimates of the effectiveness of IMF support--the β_j^{IMF} coefficients--show that the point estimates and statistical significance are sensitive to changes in the sample (Charts 1-3). Of the six recursive exercises (two for each of the three equations), four produce estimates of β_j^{IMF} that are significantly different from the whole sample estimates in at least 15 percent of the recursions. Significant deviations are more frequent in the program than in the non-program recursions. In all the recursive exercises, a sizeable share (64-100 percent) of recursion estimates are not significantly different from zero. Thus, the finding of significant effects of Fund support on growth and the debt service ratio cannot be considered robust to variations in the composition of the sample.

For the coefficients on the lagged policy variables, the recursive exercises suggest that the full sample estimates of policy effects are relatively robust (Tables 3 and 4). Typically, small proportions (in 13 of the 18 exercises, none) of the recursions produce estimates significantly different from the full sample estimates. As almost all of the full sample estimates are insignificantly different from zero, most (11) of the 18 recursions produce estimates that are also not significantly different from zero. There are a few exceptions to these generalizations, although

---

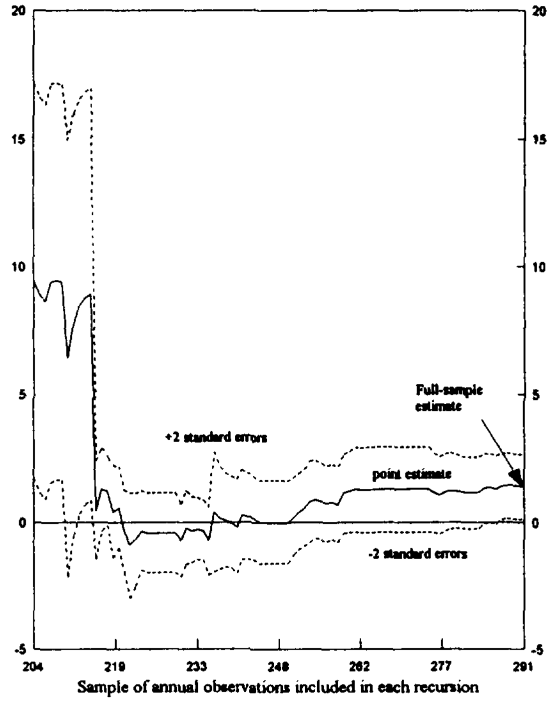
1/ Stability is assessed in terms of the point estimates and standard errors of individual parameters. An alternative approach would be to conduct Chow or Wald tests for the joint stability of the coefficients of interest on the recursive estimates of the equations. However, this was not possible because the fitted equation does not meet the requirement of both tests that the set of regressors remains constant over recursive estimates: the country dummies, which entered with statistically significant coefficients, change across the recursions.

2/ The additional program country observation was required to avoid singularity in the presence of the IMF dummy variable (d_i).

3/ These recursive procedures consider only a small subset of all subsamples that could be drawn from the data set. Thus, the range of coefficient estimates reported in Tables 3-6 does not necessarily encompass global maximum and minimum values for the sample as a whole.

Chart 1: Recursive Point Estimates and Confidence Intervals for IMF Program Dummy Variable
(Real GDP growth rates)

Program Recursions



Nonprogram Recursions

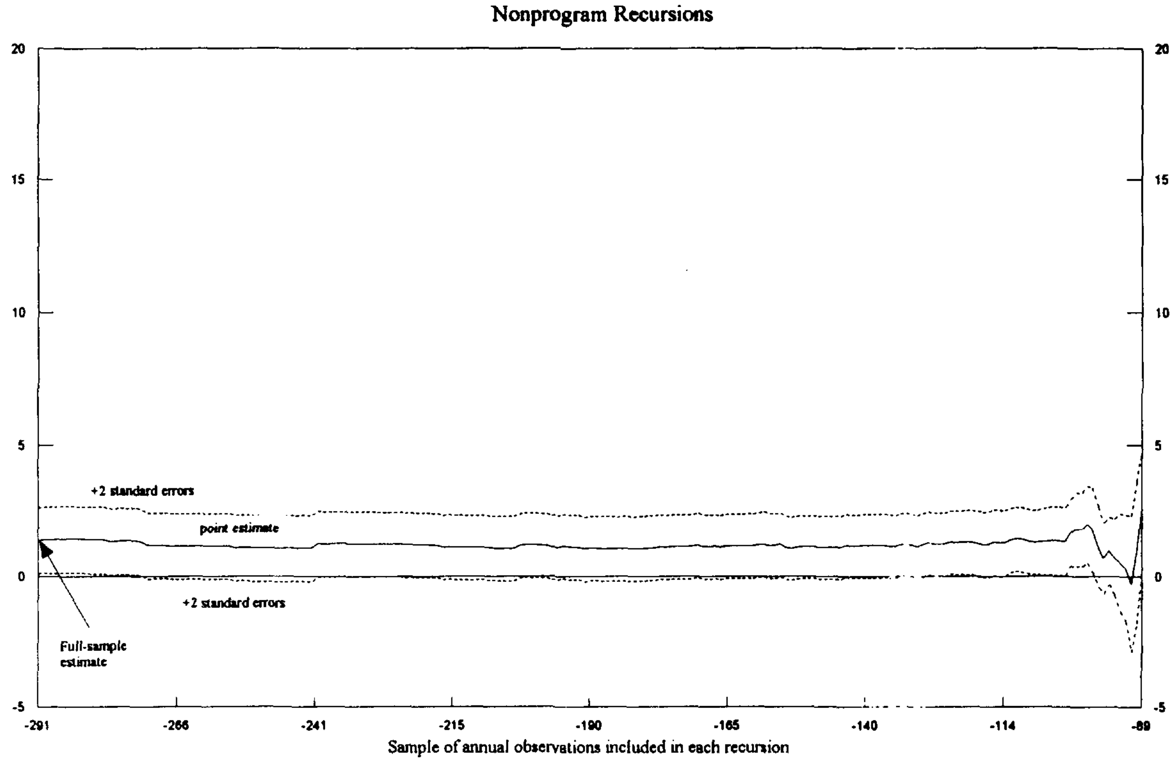
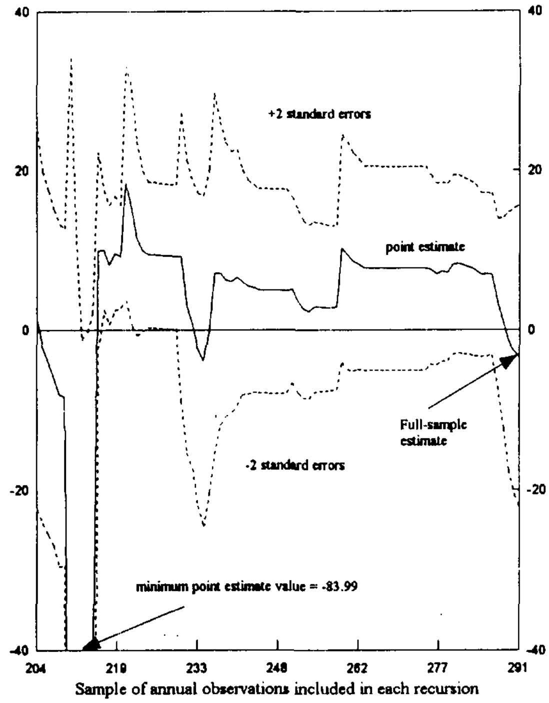


Chart 2: Recursive Point Estimates and Confidence Intervals for IMF Program Dummy Variable (Rate of inflation)

Program Recursions



Nonprogram Recursions

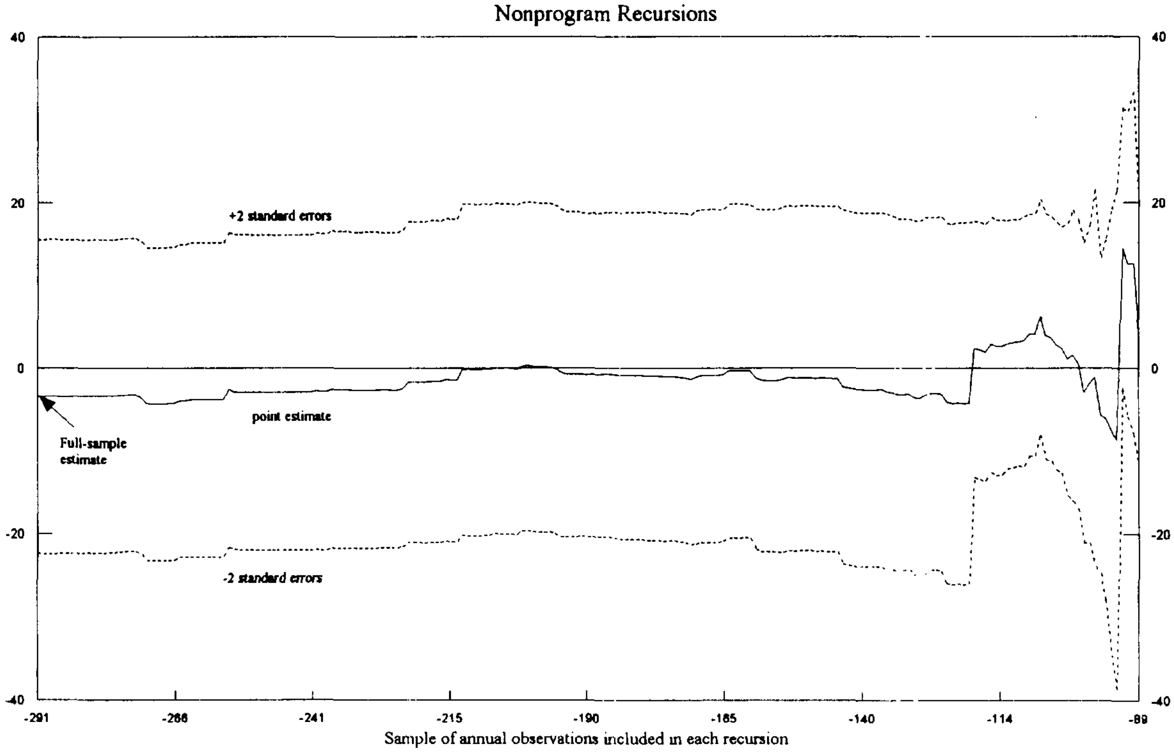
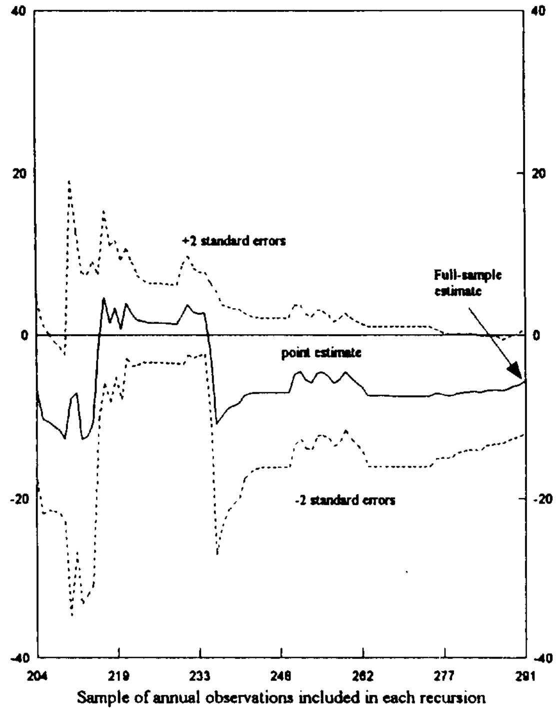


Chart 3: Recursive Point Estimates and Confidence Intervals for
IMF Program Dummy Variable
(External debt service ratio)

Program Recursions



Nonprogram Recursions

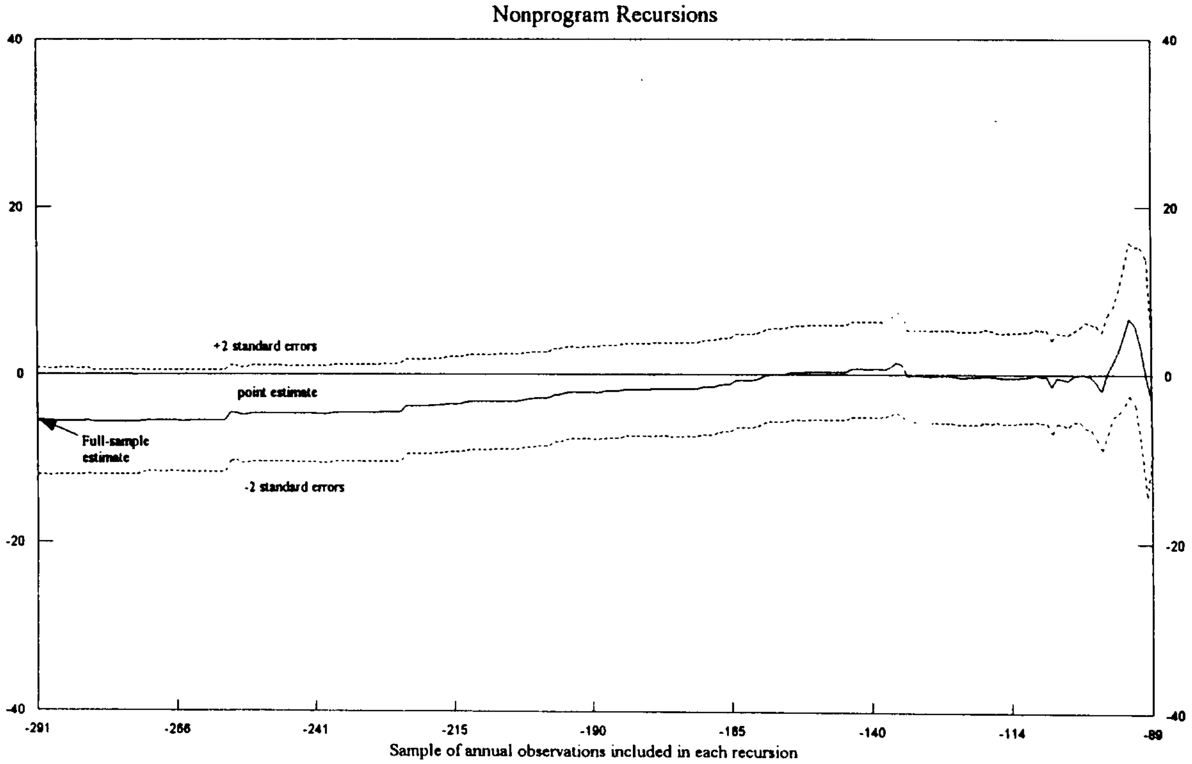


Table 3. Share of Statistically Significant t-statistics at the 5 Percent Level in Recursive Estimates of the GEE 1/

Target Variable	Lagged Fiscal Balance/GDP β_1		Lagged NDA Growth β_2		Lagged NEER Percent Change β_3	
	$H_o:\beta_1 = 0$	$H_o:\beta_1$ = full sample value	$H_o:\beta_2 = 0$	$H_o:\beta_2$ = full sample value	$H_o:\beta_3 = 0$	$H_o:\beta_3$ = full sample value
Δ Real GDP growth rate						
Share in total program year recursions 2/	0.0	0.0	0.0	0.0	0.0	0.0
Share in total nonprogram year recursions 3/	45.0	0.0	47.0	0.0	0.0	4.5
Δ Inflation rate						
Share in total program year recursions 2/	0.0	0.0	0.0	0.0	3.4	0.0
Share in total nonprogram year recursions 3/	0.0	0.0	10.9	13.9	85.6	0.0
Δ External debt service ratio						
Share in total program year recursions 2/	0.0	36.8	0.0	0.0	0.0	0.0
Share in total nonprogram year recursions 3/	54.0	49.5	57.4	0.0	0.0	2.5

1/ Excluding full sample estimates, the total number of recursive estimates was equal to 87 for program year observations, and to 202 for nonprogram year observations.

2/ Recursive procedure starts with the baseline sample (all nonprogram observations plus one program observation) and adds program observations one-by-one.

3/ Recursive procedure starts with the entire sample and subtracts out one-by-one non-program observations.

Table 4: Generalized Evaluation Estimates of Policy Parameters (β) 1/

Target Variable	Policy Variable: Lagged Fiscal Balance/GDP			Lagged Net Domestic Asset Growth			Lagged NEER (Percentage Change)		
	β_1-2 standard errors	β_1 (t-statistic)	β_1+2 standard errors	β_2-2 standard errors	β_2 (t-statistic)	β_2+2 standard errors	β_3-2 standard errors	β_3 (t-statistic)	β_3+2 standard errors
Δ Real GDP growth rate									
Full sample estimates	-0.103	-0.042 (-1.37)	0.019	0.0004	0.004 (1.82)	0.009	-0.027	-0.009 (1.03)	0.009
Recursive estimates									
•Minimum	-0.411	-0.180 (1.55)	0.051	-0.001	0.003 (1.46)	0.006	-0.055	-0.022 (-1.30)	0.012
•Maximum	-0.094	-0.033 (-1.08)	0.028	-0.004	0.013 (1.51)	0.030	-0.005	0.010 (1.32)	0.026
Δ Inflation rate									
Full sample estimates	-1.181	-0.467 (-1.31)	0.247	-0.207	-0.088 (-1.47)	0.031	0.025	0.436* (2.12)	0.846
Recursive estimates									
•Minimum	-6.101	-2.592 (-1.48)	0.917	-0.216	-0.116* (-2.31)	-0.015	-0.190	0.170 (0.95)	0.530
•Maximum	-0.954	1.358 (1.17)	3.669	-0.032	0.162 (1.67)	0.356	0.193	0.873* (2.567)	1.553
Δ External debt service ratio									
Full sample estimates	-0.157	0.097 (0.76)	0.351	-0.043	-0.020 (-1.78)	0.002	-0.052	0.058 (1.05)	0.168
Recursive estimates									
•Minimum	-0.291	-0.116 (-1.33)	0.058	-0.068	-0.041** (-2.99)	-0.013	-0.174	-0.057 (-0.97)	0.060
•Maximum	1.095	1.790 (5.149)**	2.486	-0.067	0.004 (0.11)	0.076	-0.018	0.128 (1.76)	0.275

1/ Standard errors and t-statistics computing White's heteroschedasticity-consistent variance-covariance estimator. A single asterisk indicates statistical significance at the 5 percent level; two asterisks indicate statistical significance at the 1 percent level.

none suggest the basis for questioning the general pattern of lagged policy variables having little effect on target variables.

In order to investigate the stability of the parameters in the policy reaction functions (γ_{kj}), a simplified recursive exercise over the sample of nonprogram observations was performed. The recursions began with an initial sample of 25 observations, and observations were added one-by-one. 1/ The results, reported in Tables 5 and 6, indicate that the coefficient estimates from the full sample of non-program observations are generally robust across recursions, and (except for the lagged external debt service ratio in the equation the nominal effective exchange rate policy) are insignificantly different from zero. Although the estimates of the coefficient on inflation in the exchange rate equation reveal a significant reversal of sign across recursions, sign reversals are not widespread; most of the estimates are neither significantly different from zero nor from the full sample estimate.

4. Dynamics and initial conditions

The estimation results can shed some light on the questions raised in Section II.4 about on the adequacy of the dynamic specification of the GEE. Specifically, independent estimates of γ_{kj} from the policy reaction function for nonprogram periods together with the estimates of the β_{jk} from the reduced form GEE can be used to provide an indication of the appropriateness of the static specification of the GEE--that is, that all shocks to target variables are fully reversed in one period. The product of the estimates of these two parameters for each equation is close to zero. In this case, the coefficient on the own lagged target variable in the reduced form equation ($-(\beta_{jk}\gamma_{kj}+1)$ from equation 3) would equal minus one if the assumption of complete reversion to mean were valid. Except in the growth equation, this hypothesis is rejected by the data. 2/ Nevertheless, the reduced form estimates of the coefficients on the lagged values of the target variables are significantly different from zero, suggesting that there is partial reversion to mean in the target variables: initial conditions do influence subsequent macroeconomic performances, but a richer specification of the dynamics is warranted for the sample under study.

V. Conclusions

With respect to the central objectives of this paper--to use the GEE framework to identify the independent effects of ESAF support during 1986-91 on key macroeconomic variables and to assess whether the assumptions

---

1/ The size of the initial block of observations was chosen so as to start with a reasonable number of degrees of freedom (20).

2/ The null hypothesis of an estimated coefficient equal to minus one on the lagged dependent variable is rejected at the 10 percent level for real GDP growth, 5 percent level for the inflation rate, and 1 percent level for the debt service ratio.

Table 5: Sensitivity of Estimates of Policy Reaction Function Parameters (γ) 1/

Target Variable	←Policy Variable→								
	Δ Fiscal Balance/GDP			Δ Net Domestic Asset Growth			Δ NEER (Percentage Change)		
	γ-2 standard errors	γ (t-statistic)	γ+2 standard errors	γ-2 standard errors	γ (t-statistic)	γ+2 standard errors	γ-2 standard errors	γ (t-statistic)	γ+2 standard errors
<u>Lagged GDP growth rate</u>									
Full sample estimates	-0.222	0.0024 (0.19)	0.269	-3.063	-1.090 (-1.11)	0.882	-0.795	-0.204 (-0.69)	0.388
Recursive estimates 2/									
•Minimum	-0.712	-0.159 (-0.57)	0.394	-2.793	-1.181 (-1.46)	0.431	-1.861	-0.685 (-1.16)	0.491
•Maximum	-0.205	0.156 (0.86)	0.516	-1.184	0.149 (0.22)	1.482	-0.355	0.510 (1.18)	1.375
<u>Lagged inflation rate</u>									
Full sample estimates	-0.005	0.006 (1.02)	0.016	-1.474	-0.081 (-0.12)	1.312	-0.097	-0.017 (0.29)	0.131
Recursive estimates 2/									
•Minimum	-1.196	-0.448 (-1.20)	0.300	-2.636	-1.398* (-2.26)	-0.160	-0.768	-0.439** (-2.67)	-0.110
•Maximum	-0.148	0.104 (0.82)	0.356	-0.851	0.504 (0.74)	1.861	0.192	0.511** (3.20)	0.830
<u>Lagged external debt service ratio</u>									
Full sample estimates	-0.041	-0.0007 (-0.04)	0.039	-0.584	-0.097 (-0.40)	0.389	-0.289	-0.152* (-2.22)	-0.015
Recursive estimates 2/									
•Minimum	-0.102	-0.029 (-0.79)	0.044	-1.300	-0.691* (-2.27)	-0.081	-0.708	-0.397* (-2.55)	-0.085
•Maximum	-0.127	0.075 (0.74)	0.277	-0.222	0.277 (1.11)	0.775	-0.065	0.211 (1.53)	0.487

1/ Standard errors and t-statistics computing White's heteroschedasticity-consistent variance-covariance estimator. A single asterisk indicates statistical significance at the 5 percent level; two asterisks indicate statistical significance at the 1 percent level.

2/ Recursive least squares estimates obtained by adding observations one-by-one to an initial sample of 25 nonprogram years (corresponding to a minimum of 20 degrees of freedom).

Table 6: Share of Statistically Significant t-statistics at the 5 Percent Level in Recursive Estimates of the Policy Reaction Function 1/

(In Percent)

Policy Variable	Δ Fiscal Balance/GDP		Δ Net Domestic Assets Growth		Δ NEER Percentage Change	
	Ho: $\gamma = 0$	Ho: γ = full sample value	Ho: $\gamma = 0$	Ho: γ = full sample value	Ho: $\gamma = 0$	Ho: γ = full sample value
Lagged GDP Growth Rate	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (3.4)	0.0 (0.0)	0.0 (0.6)
Lagged Inflation Rate	39.3 (57.9)	0.0 (38.8)	0.6 (3.4)	9.0 (9.6)	3.4 (7.3)	2.8 (7.3)
Lagged External Debt Service Ratio	0.0 (0.0)	0.0 (0.0)	9.0 (14.0)	0.0 (7.9)	70.8 (73.6)	5.6 (14.0)

1/ Excluding full nonprogram sample estimates, the total number of recursive estimates was equal to 178 for non-program years observations. Shares of statistically significant t-statistics at the 10 percent level are reported in parentheses.

underlying the GEE are applicable to the ESAF-eligible countries--conclusions can be summarized as follows. For output growth and the debt-service ratio, sizable beneficial effects that are statistically significantly different from zero are identified. 1/ The effects on inflation are not significantly different from zero. Diagnostic tests, however, cast doubt on the applicability of the GEE framework to the ESAF-eligible countries: the overall fit of the model is poor; estimates of the coefficients on many variables are insignificantly different from zero; regression residuals are heteroschedastic and nonnormally distributed; and the estimates of the coefficient on the dummy for ESAF support are quite sensitive to variations in the sample. A striking finding is that the counterfactual policy reaction function does not have any significant explanatory power for the sample of nonprogram observations. These results suggest that, for the sample reviewed, the GEE model is not correctly specified, parameter estimates may be biased, and conventional tests of significance may be misleading.

The results also have implications for any effort to identify the independent effects of Fund financial support. The GEE is a conceptually rigorous framework, superior to before-after and simple control group comparisons for identifying the independent effects of Fund financial support. It is, however, based on many restrictive assumptions that are necessary to define the counterfactual and to specify in a simple framework the main determinants of important endogenous macroeconomic variables. An important shortcoming of most recent applications of the GEE is their focus on the bottom line--the estimates of the effects of Fund support--with little or no evaluation of the validity of the underlying model; indeed, some studies have reported only estimated coefficients on the dummy variables and their standard errors, without diagnostic statistics or the estimates for other coefficients. One important lesson to be drawn from this study is that the validity for any given sample of the premises of the GEE methodology must be investigated before reliable conclusions about the independent effects of Fund-supported programs can be drawn from it.

This paper also points to some elaborations of the basic GEE framework that would likely help reduce its inherent restrictiveness and perhaps make it applicable in a broader range of circumstances. While this paper experimented with a few of these, such experimentation was limited by basic constraints. It is worth pointing out some of these here and indicating areas where substantially more experimentation will be needed to produce dependable estimates of the independent effects of Fund support.

- Identifying policy reaction functions and a structural economic model that are both simply specified and common to what invariably is a wide range of developing countries is difficult. Future applications of the GEE

1/ For the external debt service ratio, statistical significance was at the 10 percent level.

may need to move toward applications to single countries or a small group of relatively homogeneous countries.

- The stark distinction between periods when Fund-supported programs are in place and when they are not may not be conducive to capturing the independent effect of such programs. The Fund's influence is exerted throughout the course of negotiations, prior actions, the program itself, program extensions, and post-program monitoring. From a broader perspective, there is a question whether and how the Fund's involvement through lending arrangements should be differentiated from its influence through annual Article IV consultations (when the Fund staff provide advice on the spectrum of policies covered by ESAF arrangements).

- Many key influences on macroeconomic performance, particularly structural and institutional changes, are inherently difficult to quantify and include in the GEE framework. For ESAF-supported programs, and increasingly even for programs supported through other windows, such changes--in quantitative exchange and trade restrictions, agricultural marketing arrangements, and financial institutions, for example--are a central focus and need to be accounted for explicitly.

- The results reported in this paper suggest the presence of inertial effects in the target variables examined. In its commonly-applied form, however, the GEE assumes full and immediate reversion to mean of target variables. Also, this simple form investigates only the within year effects of Fund support on targeted macroeconomic variables. In general, however, particularly when structural and institutional changes are central to programs, effects are expected to be spread out over a far longer period. In order to enrich the dynamic specification of the GEE, longer time series than have been used to date are needed. This is a particularly onerous requirement for the ESAF-eligible countries, where the quality of the data is poor and consistent series without important breaks in definitions are not available even over the short period examined in this paper (1986-1991).

Sample Selection Bias and a Model of Program Participation

Sample selection bias in the parameter estimates of the policy reaction function (equation (2)) fitted over nonprogram observations can arise when unobservable factors that make a country more likely to seek Fund assistance also make a country more likely to have adopted a different policy package in the absence of a program than another country facing similar circumstances. To correct for this sample selection bias, following Heckman (1979) the estimated policy reaction function is augmented by an additional regressor which accounts for the bias due to nonrandom sampling.

The correction requires the estimation of a model of the probability of program status. Program status is characterized by a dummy variable d_i that equals one if the country has a Fund-supported program, a zero otherwise,

$$I_i = \delta'D + \pi_i$$

$$d_i = 0 \text{ if } I_i > I^*$$

$$d_i = 1 \text{ if } I_i \leq I^*$$

where I is a random variable, which is an index of country-specific characteristics that determines the probability of country i not having a program; and D is an n-element vector of variables that determines program status; δ is a nx1 vector of parameters; π_i is a zero mean fixed variance error term. The bias in the policy reaction function stems from a correlation between π_i and the error term (η_{ik}) in the policy equation (2). 1/ Estimates of the probability of nonprogram status were obtained using the probit model (fitted over the full sample):

$$\text{Prob } (d_i = 1) = \phi\ (-\delta'D)$$

$$\text{Prob } (d_i = 0) = 1 - \Phi\ (-\delta'D)$$

where ϕ and Φ denote the destiny and distribution function for a standard normal variable, respectively.

The estimated parameters of the probit equation are used to compute the inverse Mills ratio, i.e., the ratio $\phi(-\delta'D)/[1-\Phi(-\delta'D)]$ 2/, which is included in the policy reaction function fitted over nonprogram countries to obtain unbiased estimates of the policy response coefficients. Assuming that η_{ik} and π_i have a bivariate normal density, the inverse Mills ratio

---

1/ In terms of the reduced form GEE, a correlation between the π_i and the error term in $(\epsilon_{ij} + \beta_{jk}\eta_{ik})$ in equation (3) will lead to biassed coefficient estimates.

2/ This ratio is a monotone decreasing function (bounded by 0 and ∞) of the probability that an observation is selected into the sample of nonprogram countries, $\phi(\delta'D)$ or identically $[1-\Phi(-\delta'D)]$.

takes into account the separate effect of not having a program on policy responses. Fitted values of the estimated policy reaction function provide the counterfactual set of policies (x_{ik}) for program countries, which can be used in the estimation of the GEE (equation (1)) to obtain an unbiased estimate of β_j^{IMF}. 1/

Several explanatory variables were considered for the probit model of program status: the ratio of the overall balance of payments to exports 2/; the external debt service ratio; the ratio of the flow of external payments arrears to exports 3/; the real GDP growth rate; the inflation rate; the percentage change in the terms of trade; the growth rate in export markets; a dummy variable with a value of 1 for persistent arrears with the Fund (which would preclude a program), 0 otherwise; and a dummy variable with a value of 1 if a country had previously had a Fund-supported program, 0 otherwise--countries familiar with Fund program operations may be more likely to adopt a program. The explanatory power of these regressors was weak, and at times was associated with unexpected or counterintuitive signs. In part, the difficulty of explaining the nonprogram/program status of a country reflects the fact that the program periods under review are several years in duration: there can be substantial changes (improvements) in the macroeconomic variables that prompted countries to implement a Fund-supported program during the program period. Also, while "economic need" may turn a country towards adopting a Fund-supported program, variables such as external arrears may not explain the precise timing of the decision to start a program.

The best fit probit regression that was used to correct for sample selection bias in the policy reaction function is reported in Annex Table 1.

1/ An alternative procedure to obtain a consistent estimate of β^{IMF} would be to use the predicted probability of undertaking a program as an instrument for d_i in equation (3); in this case the policy reaction function need not be estimated. This procedure also corrects for the possibility that the choice of a country to have a program may depend on expectations of better performance in the target macroeconomic variables, i.e., the dummy is endogenous with respect to the dependent variable in equation (3). See World Bank (1990) and Corbo and Rojas (1992) for applications of this procedure.

2/ The overall balance of payments was measured including scheduled debt service payments and excluding exceptional financing, in order to provide a measure of underlying pressures on the external position. An alternative indicator of an external need to seek a Fund-supported program would be the level of international reserves. However, for several countries in the sample (CFA and ECCB country members) data on reserves are not available.

3/ Two versions of this variable were tried: one in which both positive and negative changes in arrears were recorded, and one in which only positive changes were included. The second version isolates the buildup of arrears, (typically prior to programs) and excludes the repayment of arrears (typically during programs).

Table 1: Probit Model of Nonprogram Status 1/

Variable	Coefficient	Partial Effects at the Means (percentage points)
Constant	0.756 (7.06)	0.25336
Lagged ratio of balance of payments to exports	-0.0002 (-0.32)	-0.00007
Lagged external debt service ratio	-0.005 (-2.95)	-0.00181
Lagged change in terms of trade	0.007 (1.46)	0.00249
Log likelihood	-172.041	
Pseudo-R^2 2/	0.035	
Percent correct predictions 3/	0.70	
Number of observations	291	
Jarque-Bera test for normality of residuals in auxiliary regression	47.29**	

1/ The figures in parentheses are t-statistics.
2/ The pseudo-R^2 measure is equal to $1-(\log L_{UR}/\log L_R)$, where $\log L_{UR}$ is the maximum of the likelihood function when maximized with respect to all parameters and $\log L_R$ is the maximum when maximized with respect to the constant term only.
3/ A predicted probability greater than or equal to 0.50 is associated with nonprogram status.

It should be noted that the residuals of the probit model exhibited evidence of nonnormality, in which case the equation estimates and the calculated inverse Mills ratio are likely to be inconsistent (see Greene (1993)). 1/ Therefore, the correction for sample selection bias, which relies on the consistency of the inverse Mills ratio, should be interpreted with caution.

1/ The probit residuals (π_i) were regressed upon a constant, and the residuals of this auxiliary regression revealed evidence of nonnormality; the Jarque-Bera test for nonnormality had a test statistic of 47.29, statistically significant at the 1 percent level.

Table 1. Sample of Countries and Program/Nonprogram Years 1/

Nonprogram Countries	Nonprogram Years
Afghanistan	1986/87-1988/89
Benin	1986-1988
Bhutan	1986-1991
Burkina Faso	1986-1991
Cape Verde	1986-1991
Central African Republic	1991
Chad	1986, 1991
China	1988-1991
Comoros	1986-1990
Dominica	1990-1991
Dominican Republic	1986-1990
Egypt	1989/90
Equatorial Guinea	1987-88, 1990-91
Ethiopia	1986-1991
Grenada	1986-1991
Guinea Bissau	1986, 1988, 1991
Haiti	1985/86, 1987/88-1988/89, 1990/91
Honduras	1986-1989
India	1986/87-1989/90
Kiribati	1986-1991
Lao, PDR	1986-1988
Maldives	1986-1991
Mali	1987, 1991
Myanmar	1986-1991
Nepal	1990/91
Nigeria	1986, 1988, 1990
Pakistan	1986/87-1987/88
Rwanda	1986-1990
St. Kitts and Nevis	1986-1991
St. Lucia	1986/87-1991/92
St. Vincent	1986-1991
Sierra Leone	1987/88-1990/91
Solomon Islands	1986-1991
Sudan	1986-1991
Tonga	1986-1991
Vanuatu	1986-1991
Viet Nam	1986-1991
Western Samoa	1986-1991
Yemen, AR	1986-1988
Yemen, PDR	1986-1989
Zambia	1986-1991
Zimbabwe	1986-1991

Number of countries: 42
Number of annual nonprogram observations 2/: 183

SAF/ESAF Program Countries	Program Years	Nonprogram Years
Bangladesh	1986/87-1991/92	
Bolivia	1987-1991	
Burundi	1986-1989, 1991	1990
Gambia, The	1986/87-1990/91	1991/92
Ghana	1987-1991	
Guinea	1987-1991	1986
Guyana	1990-1991	1986-1989
Kenya	1988-1991	1986-1987
Lesotho	1988/89-1991/92	1986/87-1987/88
Madagascar	1987-1991	
Malawi	1988-1991	1986-1987
Mauritania	1986-1990	1991
Mozambique	1987-1991	1986
Niger	1987-1991	
Senegal	1986/87-1991/92	
Sri Lanka	1988-1991	1986-1987
Tanzania	1987/88-1991/92	1986/87
Togo	1986-1990	1991
Uganda	1987/88-1991/92	1986/87

Number of countries: 19
Number of annual program observations: 88
Number of annual nonprogram observations 2/: 20

1/ Fiscal years t/t + 1 were considered to correspond to calendar year t if the fiscal year started before or on July 1st.
2/ Includes observations in which a SAF, ESAF, SBA or EFF arrangement was not in place.

References

Breusch, T. and A. Pagan, "A Simple Test for Heteroschedasticity and Random Coefficient Variation," Econometrica, Vol. 47 (1979) pp. 1287-94.

Corbo, Vittorio and Patricia Rojas, "World Bank Supported Adjustment Programs: Country Performance and Effectiveness", in Adjustment Lending Revisited, Policies to Restore Growth, eds. Corbo, Vittorio, Stanley Fischer, and Steven B. Webb (Washington: The World Bank, 1992).

Faini, Riccardo, Jaime de Melo, Abdelhak Senhadji and Julie Stanton, "Growth-Oriented Adjustment Programs: A Statistical Analysis", World Development, Vol. 19 (1991), No. 8, pp. 957-67.

Goldstein, Morris and Peter J. Montiel, "Evaluating Fund Stabilization Programs with Multicountry Data: Some Methodological Pitfalls," Staff Papers, International Monetary Fund (Washington), Vol. 33 (June 1986) pp. 304-44.

Greene, Joshua, "The Effects of Fund-Supported Adjustment Programs in African Countries, 1973-86," IMF Working Paper No. WP/89/38 (Washington: International Monetary Fund, 1989).

Greene, William H. Econometric Analysis (2nd Edition), New York: Macmillan, 1993

Harvey, Andrew C., Time Series Models, New York: John Wiley, 1981.

Heckman, James, "Sample Selection Bias as a Specification Error," Econometrica, Vol. 47, No. 1 (January 1979) pp. 153-61.

Hsiao, Cheng, Analysis of Panel Data, New York: Cambridge University Press, 1986.

Jarque, Carlos, M. and Anil K. Bera, "A Test for Normality of Observations and Regression Residuals," International Statistical Review, Vol. 55, No. 2 (1987), pp. 163-72.

Khan, Mohsin S., "The Macroeconomic Effects of Fund-Supported Adjustment Programs,", Staff Papers, International Monetary Fund (Washington) Vol. 37 (June 1990) pp. 195-231.

Killick, Tony, "Can the IMF Help Low-Income Countries?", The World Economy, forthcoming.

Lucas, Robert E., "Econometric Policy Evaluation: A Critique," pp. 19-46 in K. Brunner and A. H. Metzer (eds.), The Phillips Curve and Labor Markets, Amsterdam: North Holland, 1976.

Maddala, G., Introduction to Econometrics, 2nd Edition, New York: MacMillan, 1992.

Priestley, M.B., Spectral Analysis and Time Series, London: Academic Press, 1981.

Schadler, Susan, Franek Rozwadowski, Siddharth Tiwari and David O. Robinson, Economic Adjustment in Low-Income Countries - Experience Under the Enhanced Structural Adjustment Facility, IMF Occasional Paper No. 108 (Washington: International Monetary Fund, 1993).

Swamy, P.A.V.B., "Efficient Inference in a Random Coefficient Regression Model", Econometrica, Vol. 38, No. 2 (March 1970), pp. 311-23.

Turnovsky, Stephen J., Macroeconomic Analysis and Stabilization Policy, London: Cambridge University Press, 1977.

White, Halbert "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica, Vol. 48, (1980), pp. 817-38.

World Bank, Adjustment Lending Policies for Sustainable Growth, Policy and Research Series, Paper 14, (Washington, 1990).