CENTRE D'ÉTUDES
ET DE RECHERCHES
SUR LE DÉVELOPPEMENT
INTERNATIONAL

DOCUMENT DE TRAVAIL DE LA SERIE

ETUDES ET DOCUMENTS

E 2003.06

Performance-Based Conditionality: A European Perspective[*]

Christopher ADAM[⊗], Gérard CHAMBAS[⊗⊗]
Patrick GUILLAUMONT[⊗⊗], Sylviane GUILLAUMONT-JEANNENEY[⊗⊗]
Jan Willem GUNNING[⊗⊗⊗]

Original Version: 18 December 2002
Second Version: 15 February 2003
This version: 10 March 2003

28 p.

[⊗] University of Oxford.
[⊗⊗] CERDI, CNRS Université d'Auvergne
Corresponding author Patrick Guillaumont (P.guillaumont@u-clermont1.fr)
[⊗⊗⊗] Free University, Amsterdam

* We gratefully acknowledge the contribution of the European Commission to the research on which this paper draws. We also wish to record our appreciation to the many officials and representatives of donor organization we met in the course of our four case studies. The views expressed in the paper are, however, solely those of the authors and cannot be taken to be the official position of the Commission, donor agencies or recipient governments.

Abstract

For several years, the conditionality underpinning budgetary support to developing countries has been the object of severe criticism. This criticism has led to the belief that the “ownership of policies” by the recipient country governments is essential for the effective implementation of economic reform programmes. To this end, it has been suggested that traditional conditionality based on policy measures be replaced by a conditionality in which aid is based on performance outcomes. In this spirit, the European Commission has reformed its adjustment aid. The present article provides an assessment of this experiment based on the experiences in Benin, Burkina Faso, Madagascar and Uganda. It demonstrates the danger of a shift towards intermediate indicators, which are too distant from the final impact of the policies, and suggests several improvements to implement genuine performance-based conditionality.

Key words : Aid, fiscal policy, developing countries, conditionality

Résumé

Pendant de nombreuses années, la conditionnalité des appuis budgétaires aux pays en développement a été l’objet de fortes critiques. Cette critique a inciter à considérer «l’appropriation » des politiques comme une condition essentielle à la mise en œuvre effective des programmes de réformes économiques. A cette fin, il a été suggéré de remplacer la conditionnalité traditionnelle fondée sur l’adoption de mesures de politique économique par une conditionnalité où le versement de l’aide est subordonné à des résultats. En suivant cette orientation, la Commission Européenne a réformé son aide à l’ajustement en a. Cet article fournit une évaluation de cette expérience à partir des cas du Bénin, du Burkina Faso, de Madagascar et de l’Ouganda. Il met en lumière le danger de glisser vers une conditionnalité fondée sur des indicateurs intermédiaires, trop éloignés de l’impact final des politiques. Il suggère différentes améliorations pour mettre en œuvre une conditionnalité effectivement fondée sur la performance.

Mots-clés : Aide, politique budgétaire, pays en développement, conditionnalité

Introduction

In recent years, the European Commission has embarked on a radical reform of the conditionality underpinning its budgetary support to ACP countries. Since 1999, financing conventions signed with these countries have included a "variable financing tranche" whereby the disbursement of aid is no longer conditioned on the implementation of specific policy measures or actions, but rather on outcomes in key economic and social sectors.

This European reform of conditionality has occurred against the background of a broader international debate on aid effectiveness from which has emerged a consensus in a number of areas, but residual tensions in others. For example, there is a widespread acceptance that, at least in the presence of recipient governments committed to good economic policies, untied aid appears better able to support a consistent programme of public expenditure management than aid allocated to targeted spending initiatives or to projects. Nonetheless, the international community remains of the view that it would be inappropriate to relinquish conditionality altogether, regardless of whether aid was in the form of loans or grants. And while there was agreement on the need for the reform of conditionality, there is less consensus over how best, and to what extent to replace traditional forms of conditionality. The European Commission contributed significantly to this latter debate[1].

Aside from the most radical element of the critique of conditionality which viewed donor-supported reforms as poorly adapted to the specific situation of each developing country, three recurrent themes shaped the critique. The first was the concern that the formalism and multiplicity of conditions, combined with a lack of coordination between donors, generated a conditionality that was discordant, lacked credibility, and encouraged recipient governments to accept negotiated conditions in order to obtain money, without having either the conviction that the agreed economic policy reforms were useful, or having any real intention of implementing them[2]. Second, the "all or nothing" character of aid payments which emerged

[1] The European Commission's promotion of a reform of conditionality proceeded in several stages: interventions in the framework of meetings of the Special Programme for Africa (SPA); the commissioning for academic work; preparation of a document on the reformulation of conditionality (October 1995); developing a Communication on the issue to the Council of Ministers; and publication of a Communication from the Commission to the Council and to the European Parliament *: Community support for economic reform programmes and structural adjustment: review and prospects (2000).*

[2] Mussa and Savastano (1999), for example, note that between 1977 and 1997, almost half of IMF programmes were disbursed at less than 75% of the loan.

from this system generated discontinuities in and raised the uncertainty of aid flows, further compromising the successful implementation of reforms.[3] The third, and in many senses the overarching element in this critique, was that traditional conditionality undermined the aid recipient's claim to 'ownership' of policy reforms.

All multilateral and bilateral aid institutions accepted the idea that the efficiency of budgetary aid requires ownership by the recipient countries. Ownership operates at a number of levels. First, as is shown in the literature on agency theory, strong domestic ownership can work in the interest of both donors and recipients (Tirole 2001, Khan and Sharma 2001). Since the principal (here the donor) must rely on the government of the recipient country (the agent) to pursue certain objectives, and that the agent has the possibility of following his own interests rather than that of the principal, the latter finds every advantage in the agent having the same objectives as him(Adam and O'Connell, 1999). Second, and in the same setting, strong ownership may provide for quick and efficient response from the government to unexpected events and problems not identified when the economic policy programmes are drawn up. Most fundamentally, however, ownership of reforms is indispensable if local democratic accountability is not to become devoid of meaning in the recipient country and if the legitimacy of its government is not to be called into question.

Although the objective of policy ownership is well-defined, the means of achieving it remain uncertain. Some authors have argued for a greater level of selectivity among recipient countries in favour of countries committed to policy objectives favoured by donors; this is the spirit of the IMF reforms recommended by the Meltzer Commission requested by the US Congress, according to which the Fund's role as lender-of-last-resort should be reserved for "well-managed" emerging countries.[4] Others have recommended the use of "floating tranches", according to which recipient countries retain the ability to choose the date for the implementation of reforms which are likely to trigger aid disbursements. This practice effectively gives the recipient a degree of freedom but also moves the donors away from having to confront the dilemma of either brutally interrupting the aid programme, or waiving the condition not satisfied at the risk of removing all credibility from the sanctions. It has

[3] Several of the authors of the present study contributed to this criticism. See notably: Guillaumont . and Guillaumont-Jeanneney (1994, 1995, 1995a) in conjunction with the European Commission); Collier *et al*. (1997); Collier. and Gunning (1999); Gunning (2000).
[4] Meltzer *et al*. (2000).

been used in Africa by the World Bank since 1995 in the framework of "High Impact Adjustment Lending" and is considered with interest by the IMF.

A third approach, recommended and implemented by the European Commission, has been to substitute instrument conditionality with performance-based conditionality. In this case, adjustment aid is no longer tied to the adoption of particular policy measures but to the performance achieved by the countries in specific agreed economic or social domains. Underpinning this recommendation, is the notion that performance conditionality allows for better ownership of reforms, since the choice of instruments would reside with the country; it avoids arbitrary judgement on multiple heterogeneous economic policy measures; it facilitates gradual and progressive support according to the degree of progress of performance relative to outturns; and by eliminating the scope for discordant conditionality, it supports better coordination between donors.

Europe has moved ahead rapidly to implement these ideas.[5] This paper is concerned with the design and initial impact of the European Commission's reforms. By the end of 2001, 28 European financing arrangements with the ACP countries included some element of performance-based conditionality. However, the gradual nature of the reform in all these countries meant that in some cases traditional conditionality (based on policy measures) was maintained alongside the new conditionality and that in a number of programmes, the new conditionality sometimes only required the *definition* of the outcomes indicators rather than their evaluation (see Table 1). Our review therefore draws principally on experiences in four countries where the new conditionality has been applied most completely, although even here, the experiment is still very much in progress. The four are Benin, Burkina Faso, Madagascar and Uganda.[6]

The next section of the paper sets the scene and provides a brief sketch of the programme design in each country. The remainder of the paper is concerned with two aspects of the

[5] The World Bank and IMF have had reservations about introducing a greater degree of performance-based conditionality into their lending arrangements. However, a recent IMF document, *The Modalities of Conditionality – Further Considerations*, Policy Development and Review Department, January 8, 2002, provides a favourable presentation of outcomes-based conditionality and payments according to a floating tranche. For their part, Lerrick and Meltzer (2002) have suggested that the IDA should accord grants to low-income countries and releases its aid according to social results (health, education).

[6] This examination was conducted at the request of the European Commission between March 2001 and June 2002.

reform. In Section three we consider the extent to which the new practice meets the objectives which motivated it, namely the promotion of greater ownership of reforms by recipient countries, improved coordination between donors, and greater consistency and predictability in resource flows. In Section four we focus on the operational question of how performance indicators should be identified and performance assessed.

1. New European conditionality in four countries: Issues in Programme Design

Madagascar was the first country is which a new type of programme was adopted. The programme was rapidly developed and underwent substantial modification during its implementation. By contrast, in Burkina-Faso the ground for a move towards performance-based conditionality was prepared at length by the so-called "conditionality test"[7], which took place from 1997 to 2000. The test emerged from a sequence of five meetings or "missions" in which the donors and Burkinabe authorities developed a set of outcome indicators designed to reflect policy implementation and defined the modalities by which these indicators could serve as budgetary aid payment criteria. Following this, two "new formula" support programmes had already been implemented and evaluated in this country. In Benin, the new European conditionality was developed within the context of a broader World Bank supported reform programme for the entire public expenditure system. The World Bank chose Benin (and Jordan) to launch their own budgetary reform based on the development of expenditure programmes for major ministries incorporating explicitly defined objectives and outcome indicators. Finally, Uganda is widely seen as a standard-bearer in the process of reforming economic policy dialogue and is considered as having internalised the reform process. As a consequence, outcomes indicators have been part of the policy dialogue for some time, although not as explicit instruments for aid disbursement. The new European procedure is less advanced in Uganda than in the other three countries, with only one programme having been implemented which had still not been evaluated at the time of writing.[8]

In each country performance indicators were chosen following dialogue between the countries and the European Commission. This ensured that they were consistent with initiatives taken either by the government or other donors. Thus in Benin, the indicators were drawn from the

[7] Burkina-Faso had already been chosen as early as 1996 by the Special Programme for Africa (SPA) as a test case for the new approach to conditionality.

budget programme and were selected during negotiations between the authorities in Benin and the European Union, while in Madagascar, the determination of indicators was part of the conditionality in the very first tranche of the program. In Burkina-Faso, the indicators were determined based on the "conditionality test" and the *Poverty Reduction Strategy Program* (PRSP), which includes a list of priority indicators. A similar process occurred in Uganda, where the indicators came from the *Poverty Eradication Action Plan* (PEAP), the forerunner of the PRSP.

Once indicators have been selected, the new procedure then relates aid payments to actual performance relative to target levels for these indicators Disbursement is graduated according to a three-level scale of (0 percent; 50 percent, and 100 percent) depending on the level of performance, thereby avoiding the "all-or-nothing" feature of traditional conditionality.

Practice varies across countries in terms of the number and domain of the indicators (see Table 3). For Benin, Burkina-Faso and Uganda programmes were based on between 9 and 15 conditions. Indicators concern budget administration, education and health sectors, except in the first Benin programme where disbursement was linked to indicators concerning the health sector only. In contrast in Madagascar disbursement was initially tied to no fewer than fifty-four indicators, but similarly concerned with public finance administration, with education and with health. The number of indicators adopted later fell to thirty-four . Despite the variation across countries, programmes shared many common indicators. For example, all countries' programmes included measures of the extent of the decentralization of budget expenditures; measures of implementation rates for priority sectors; utilization rates for health facilities; and vaccination coverage rates.

The weighting given to each sector (public finance administration, education, health) in determining overall disbursement also varies. In Benin, Madagascar and Uganda performance in each is equally weighted (although this is implicit in Uganda where no explicit weights are provided). Only in Burkina-Faso has there been an attempt to discriminate between sectors.[9]

---

[8] Table 2 in the appendix summarises the state of progress of the programmes during our missions in the four countries. The case of Uganda is discussed in greater detail by Adam and Gunning (2002).

[9] Thus public finance indicators were given a weight of 50 percent in the two successive programmes, while the weight on education indicators increased from 15 percent in the first programme to 25 in the second, and that on the health sector fell correspondingly from 35 percent to 25 percent. Within each sector, however, all the sub-indicators were equally weighted.

Importantly, the graduated design of the new conditionality meant that from the start disbursement was modulated according to performance. In Benin, for example, the EU's assessment of performance against the targets led to 67 percent of the variable tranche of the first Benin programme to be disbursed. In Burkina the 79 and 60 percent of the variable tranche of the two Burkina programmes were disbursed in the light of performance respectively, and in the Madagascar programme the rate was 71 percent. Evaluation of the Ugandan programme is planned for November 2002. However, since the performance-related 'variable tranche' itself accounted for a relatively small proportion of the total aid flow from Europe, the volume of aid withheld through this mechanism was not substantial. In Burkina-Faso, for example, the 21 percent of the first variable tranche which was withheld accounted for about 4.2 % of European budgetary aid, which itself represented just under half of total European aid in 2000: the 'retention' rate under the variable tranche thus represented about 2 % of total European aid in 2000 (which itself accounted for only a fraction of total aid to Burkina). Similar calculations for other programmes reveal similarly weak 'withholding effects'.

3. New conditionality and its objectives

The new European practice had three main objectives: stronger domestic ownership of the reforms, better coordination between donors and reduced uncertainty surrounding donor aid flows. In this section we consider how close actual practice in the four countries has come to meeting these objectives.

3.1. Reform and Policy Ownership

The centrality of the role of domestic ownership of economic reform was reinforced during the nineties by the progress of democracy in many aid-receiving countries. In autocratic regimes the interference constituted by direct policy conditionality, often aimed at eroding resistance to reform by established interest groups could, in the eyes of the donors, appear, rightly or wrongly, as a valid defence of the interests of the populations. When regimes are more democratically accountable, however, this interference is much more difficult to defend since it risks depriving the domestic authorities of their legitimate freedom of choice.

However, although domestic ownership is all the more necessary as countries become more democratic, democratic accountability does not automatically create the conditions for an efficient ownership. Efficiency in this context requires a capacity to design and implement policy. All four countries studied here are defined by the UN as least developed countries, where the managerial capacity of the public sector is perceived to be limited. Not surprisingly, there was no substantial shift in ownership with the recent implementation of its new conditionality by the European Commission; even if the capacity had been there it is still doubtful there would have been a profound change since the new conditionality is limited to a still modest proportion of the total European support received by the country.

A culture of results

The Commission's attempt to promote ownership by modulating aid flows according to performance rests on two ideas. The first is that the ownership implies that the political leaders are concerned about the results of their action: granting aid according to the results is thus a way of promoting attention to the outcomes of policy choices. The second idea is that if the aid is allocated according to results rather than simply the measures taken by the countries, governments having *de facto* greater freedom in their choice of measures (and their timing) will implement them more resolutely.

The new European practice has experienced greater success in promoting a "culture of results" than in actually shifting responsibility for policy formulation to recipient governments. The evidence of a cultural shift of ideas can be seen in many areas. For example, in Burkina Faso the preparation of the "conditionality test" has contributed to this shift as witnessed by the involvement of the National Institute for Statistics and Demography, both in conducting surveys and opinion polls, especially with the parents of pupils and health centre patients, which influenced the design of the variable tranche of European aid, and in creating a National Poverty Observatory. The "conditionality test" also played an important role in the preparation of the PRSP with the matrix of indicators adopted by the PRSP corresponding exactly to the conclusions of the test.

It is, of course, difficult to isolate the influence of the European initiative from that of the other concomitant factors on the improvement of the statistics and monitoring of the outcomes of the policies undertaken. In part this reflects a range of other initiatives launched concurrently with the new conditionality. In Madagascar, for example, even before the implementation of the European programme, a considerable externally-supported project of improving statistical knowledge had been undertaken and naturally served to collect the information required by the European budgetary aid programme. Similarly in Uganda, a considerable effort towards the definition, collection and diffusion of statistics relating to the social sectors was in place before the introduction of the new European procedure, while in Benin, the effort towards collection and monitoring of the outcomes indicators required by the European initiative was dominated by the experience of budgetary practice reforms initiated by the World Bank.

Anecdotal evidence suggests some positive response. For example, when evaluations highlighted low utilization rates for rural health centres in Burkina-Faso it was decided to postpone the construction of new centres in order to re-equip and improve service provision in the existing ones (medicines, material, personnel). Similarly, it was judged more appropriate to improve the functioning of poorly frequented existing schools than to increase the number of schools built.

Policy Accountability

Ownership also requires governments to exercise their responsibility, both in terms of the decisions taken and their implementation. As with the culture of results it is difficult to differentiate between the impact of the new European conditionality and that of the simultaneous changes in the practices of the other donors. Indeed, the Highly Indebted Poor Countries (HIPC) Initiative is founded on the belief that the efficient use of the proceeds from debt relief to alleviate poverty can only be successful if there exists strong domestic ownership of the programme. HIPC governments are thus required, as a condition of debt cancellation, to establish a poverty reduction strategy themselves and to include the representatives of the “civil society”. Whatever the case, whether the new European conditionality alters the sense of accountability depends primarily on three categories of factors: the nature of the indicators; governmental stability and coordination; and the credibility of the sanction respectively.

First, the reality is that the new conditionality has not yet created the space for governments to exercise discretion over policy choices. A large number of the indicators which served as payment conditions are still concerned with specific actions to be taken, rather than actual outcomes: this is certainly the case for indicators relating to budgetary management, but it also applies in the health and education sectors, for example the indicators relating to staffing levels and provision of equipment. This incompatibility with the intentions of the new conditionality reflects in part the difficulty in collecting genuine indicators of the final outcomes for the policies. It has also been encouraged by a certain distrust with respect to the desire or capacity of the national leaders to conduct their own education and health policies efficiently – itself reflecting a deeper conflict about ownership on the one hand and the alignment of the interests of recipients and donors on the other. The observed retreat from performance indicators is not specific to the new European procedure. In the case of Uganda, where government is in the "driver's seat" and where systems of allocating donor funds to the health sector on the basis of performance indicators existed prior to the redesign of European aid, a similar tendency can be noted for the donors to request particular measures and to return to the very practice that results-based conditionality was designed to avoid. We return to the difficult but central issue of the nature of the indicators in the next section.

The experience of Burkina Faso has highlighted the fact that the ownership of policies, notably in the social domain, requires that a key minister (Prime Minister or Finance Minister) takes direct responsibility for creating cohesion within the entire government team, responding to the multiplicity of the factors which influence the evolution of the social indicators, and coordinating action. Similarly in Uganda, ownership was only given genuine meaning at critical points in the reform process when individuals – occasionally the President himself – were seen to take personal charge of decision making. Clearly, when responsibility is diffused, or where turnover of ministers in charge is rapid, actions become disconnected from their consequences and the incentive to monitor diminishes.

The third important factor of efficiency of the ownership procedure is the *credibility* of financial sanctions in the event of poor performance. This problem of credibility occurs, of course, for all conditional aid and is at the heart of the literature on conditionality.
For the sanction to be credible, it must first be widely known. In Benin it was the evaluation of the first programme containing a variable tranche (and the withholding of part of the tranche) which made political leaders aware of the financial stake of obtaining the programmed results. When a certain retention rate has already been applied to the variable tranche, the sanction of the new conditionality automatically becomes more credible. Second, the sanction must be commensurate with performance (or lack of it). The "all-or-nothing" form of payment underpinning traditional conditionality clearly violates this requirement; one

reason why European conditionality is more likely to be credible is its progressive character. The graduated structure of payments based on a number of indicators allows payment according to the relative level of performance and avoids the binary choice between an appraisal of success or failure. Clearly, this can not be effective if the financial stake is too weak. As noted, the small size of the European support to the variable tranche (approximately 2% of total aid) goes some way to explaining the lack of apparent attention initially paid to European conditionality. Finally, a sanction has to be credible in the sense that donors will, indeed, carry through with the threat if performance targets are not met. Here the experience of Uganda demonstrates well the problem of implementing sanctions when there is doubt about the validity of targets. The education sector review of 2001 indicated significant performance short-falls but the response of the donors was to question the targets rather than the performance and, indeed, in some instances to *increase* disbursement in order to improve the target-setting process itself.

3.2. Donor Coordination

Provided donors agree on the set of targets and the criteria for success, result-based conditionality is better able to avoid the problems of lack of donor coordination.

Some progress but tensions remain

The European initiative has been accompanied by an improvement in donor-coordination, supported in particular by the Special Programme for Africa (SPA), a key forum of the principal donors to Africa. It was here that the European Commissions proposals for the reform of conditionality were first presented (in 1996) and that the idea for the joint pilot experiment in Burkina-Faso -- the “conditionality test” mentioned above – was mooted. The ‘conditionality test’ brought together all the donors to Burkina-Faso, regardless of whether they provided budget support or not, allowed competing points of view to be assessed and demonstrated the utility of combined work on the objectives to be pursued and the indicators to be adopted. Despite being specific to Burkina Faso, the lessons were widely broadcast. More generally, though, the international context of budget support for Africa has changed in favour of better donor-coordination in the wake of the HIPC initiative. The requirement to develop coherent PRSPs has led to greater coordination between creditors and debtors and between creditors themselves. This coordination is reflected in the commitment by both donors and government that the PRSP should constitute the common reference for their actions. Nonetheless some serious tensions remain.[10]

In recent years, the World Bank has been a major advocate of the reform of conditionality. In particular its report *Assessing Aid* (World Bank, 1998), based on the econometric study of Burnside and Dollar (2000), argued that international aid is only truly effective in promoting growth in countries which implement good economic policy and that aid itself has little impact on policy choices. This report contributed to the emerging consensus that conditionality was not, in itself, capable of creating the necessary incentives to produce policy reform. There is, however, a profound disconnect between theory and practice within the World Bank. At an operational level, the Bank remains reluctant to link aid to economic performance explicitly and remains attached to a conditionality dealing with, at least in part, the *implementation* of economic policy measures. By contrast, in recent years, the European Commission has remained somewhat withdrawn from the public debate but has gone further than the Bank and the bilateral donors in applying the new ideas concerning conditionality. Thus, despite an agreement on the objectives of a reform of conditionality, divergences remain. The Bank's reluctant to renounce its dialogue on economic policy measures is in evidence in the four countries examined here. It provides these countries with considerable technical aid together with its financial aid (for example, Burkina Faso receives about 300 missions per year), and it does not consider that its assistance is an obstacle to the domestic ownership of policies. It is certainly true that, in its programme evaluation, the World Bank includes the results of these policies and also agrees to consider indicators common to the other donors, but to this is added an assessment of policy implementation, defined bilaterally with the country.

The Bank also eschews an indicator-by-indicator evaluation where each partial result would trigger a certain amount of aid (as in the case of the variable tranche of European aid), preferring instead a global appraisal of the policy results.[11]

[10] However, in some ACP countries, the European Commission and several European states have agreed on common evaluation of results and on a disbursement of their budgetary aid according to these results .

[11] It also argues that since the Bank's only support instruments are loans, legal restrictions limit it to maintaining a conditionality of means – which, being directly verifiable are contractible – as opposed to outcome which are necessarily uncertain and partially independent of actions of the debtor.

An obvious concern is that the co-existence of a dual practice undermines the effectiveness of the new European conditionality to promote greater policy ownership by recipient governments. It is not immediately clear that this concern is too severe since the combination of support, some conditioned by a dialogue on policies, some by a dialogue on results, may well be complementary, with the Bank seeking to influence economic policy choices, and the European Commission requiring governments to pay permanent attention to the results of these polices. To a certain extent, this resolves the problem of coordination, albeit in an imbalanced manner such that in the limit the division can result in the evaluation of results by the Commission becoming an evaluation of the policies recommended by the World Bank. It is also true that the national government then finds that it has an increased power of negotiation, for it discusses policies with one partner exploiting the expected results on which it will be judged by the other, while with the second partner it evaluates the results with regard to policies for which the first partner is, in part, responsible. Thus, if the procedure of the European Commission contributes to making the governments more responsible in implementing their economic policy, it should also contribute to improving the quality of the dialogue these governments hold with the Bretton-Woods Institutions.

.

3.3. <u>The predictability of donor support</u>

One of the criticisms of traditional conditionality was the extreme variability of the resulting payments; depending on the success or failure of the entire programme, payments were either made in full or withheld in their entirety. Often this irregularity of aid flows contributed to macroeconomic instability, while its unpredictability often compromised the introduction of potentially costly (and partially irreversible) reforms for fear that they would not be fully funded for the duration (reform of the tax system, civil service form, employment of new teachers, etc.). In principle, the design of the new European system avoids the risk of “all or nothing” disbursement. However, as we noted above the new procedure applies only to a relatively small proportion of total European aid (and hence an even small fraction of total aid), and has only been recently implemented; it is therefore not yet possible to ascertain what impact, if any, this has had on the predictability of aid flows. Moreover, the specific contribution of the European mechanism to smother aid flows is also hard to determine if, as is the case, other donors are simultaneously undertaking (discretionary) performance appraisals.

The desire on the part of the donors to smooth aid flows and make them more predictable must still be traded-off against their ability to threaten (and use) the suspension of aid to enforce their interests, a fact which is the basis of all conditionality. The uncertainty regarding the sum of payments should nevertheless diminish with the acquired experience of an outcomes-based conditionality. It is mainly in relation to the objective of ownership that we now examine the relevance of the indicators adopted.

4. Measuring Outcomes

Since the new conditionality is supposed to rely on results indicators, the choice of these and the manner in which their evolution is evaluated are essential to the success of the new procedure. The choice does not only relate to the technical quality of the indicators; it should also reflect the priority objectives of economic policy. And the combination of indicators should facilitate the formulation of an overall judgement for the policy. In the light of this, we consider three dimensions of this choice: the coverage and quantity of indicators; the level at which the results are observed – intermediate or final; and finally, how indicators should be used to modulate the level of budgetary aid.

4.1. Coverage and quantity of indicators

The European Commission has concentrated on indicators relative to education and health, directly linked to the evolution of poverty, and has buttressed these with indicators of budgetary efficiency. Indicators of budgetary management do not really conform to the principle of performance-conditionality. Moreover, experience has shown that it is dangerous to judge the quality of budgetary management by means of a few particular indicators, such as the level of arrears or payment lags. Whatever its legitimate concern about good governance, the European Commission cannot rely in this field on robust outcome indicators and so has to leave it out of the performance conditionality and should concentrate on performance in education and health sectors

As we noted above, though most countries focussed on a relatively small number of indicators to manage the variable tranche of new European support arrangements, there were exceptions. In Madagascar, the large number of indicators (fifty-four) seems to have been desired by the government itself, who felt that progress against some of the many indicators would limit the risk of no disbursement at all, although it may also have reflected the interest of the European Commission delegation to establish a relatively complete vision of the sector considered. However, the increase in the number of indicators causes extreme complexity in the evaluation, which itself runs counter to the objective of rapid disbursement. By contrast, limiting the number of indicators has the advantage of establishing clear priorities and rejecting indicators whose meaning is too specific. Thus the choice of the number of indicators is closely linked to that of the level of observation of the results.

4.2. The level of observation

The "conditionality test" classified health and education indicators into four categories, defined according to their increasing proximity in relation to the final result: (i) indicators relating to actions undertaken by the government (*input indicators*), for example the sum spent on the construction of schools or health centres; (ii) indicators of implementation (*output indicators*), for example the number of schools or health centres opened in rural areas or the number of school books per pupil; (iii) intermediate indicators of result (*outcome indicators*), such as the rate of passage from primary to secondary education or the rate of frequentation of the health services; and finally (iv) indicators of final result (*impact indicators*), such as the rate of literacy or the rate of child mortality.

The new European conditionality was largely based on intermediate outcome and output indicators, rather than indicators of impact or final result. It is the latter which genuinely reflect the state and evolution of poverty, even if they express a specific but essential dimension of poverty (or its absence); in this case the state of health and the state of acquisition of knowledge. In other words, less interest was paid to the very levels of education and health than to the use of health (rate of utilization) or education (rate of schooling) services and the factors likely to increase their use. It seems that there were three reasons for this choice: first, these indicators seem to highlight the policy deficiencies with regard to health or education; second, they are easy to collect by the administrative route compared to indicators of impact; and finally, they are considered to be less dependent on exogenous factors, outside the influence of public action. However, the use of these intermediate indicators runs counter to the promotion of ownership of policies.

Some civil servants interviewed in Burkina Faso have underline that, since the intermediate indicators are close to economic policy instruments, the new conditionality is not very far from the traditional conditionality. Furthermore, the state of health and education depends on a multitude of actions implemented outside the corresponding sector. For example, the rate of child mortality can be reduced not only by improved access to health services, but also by other actions favouring cleansing (drinking water, fight against insects and parasites), as well as by increases income or improvements in education. Thus, the further the indicators are from the final result, the less faithful is the representation they give of the overall government policy. Moreover, intermediate indicators, insofar as they directly reflect the action of certain actors, lend themselves more easily than impact indicators to manipulations. Finally exogeneous factors have an influence on the intermediate indicators as well as on the final impact indicators. Consider, for example, the rate of utilization of public health centres. Their increase is supposed to reflect an improvement in the quality of the services provided, thus contributing to better health, but it can also simply occur in the event of epidemic or an increase in any other form of morbidity or in the event of the closure of a private health centre. Not only do the intermediate indicators not escape the influence of exogenous factors, but the direction in which they are influenced is ambiguous.

The validity of impact indicators

Hence there is a case to shift more towards final result indicators. These figure in each country's PRSP and are part of the list of "millennium development goals" adopted by the international community. We exclude from the discussion indicators of income poverty, which are very difficult to collect and always subject to discussion on the measurement of the poverty line. On the other hand, in the health and education sectors, it is possible to identify a small number of impact indicators likely to be collected without excessive difficulty and which enjoy a universal meaning.

In the field of health, the most pertinent indicator for low-income economies is incontestably the rate of child mortality. These data are currently supplied on a regular and reliable basis, but with a frequency of three to five years, by the DHS (Demographic and Health Survey)[12]. This indicator is preferred to life expectancy at birth due to the absence of equivalent information for the mortality quotients at the different stages of life. However, it is seen to

[12] The Uganda report notes that this indicator curiously does not figure in the HSSP (Health Sector Strategic Plan 2001-2004).

reflect relatively well the level of health of the whole population including, to a certain extent, the impact of AIDS through the transmission of the infection from mothers to new-born children. It is an indicator which is relatively variable in the short term as it relates to the most vulnerable part of the population. For example, for Madagascar from 1997 to 2000, it falls from 159 ‰ to 142 ‰, whereas in Burkina-Faso it rises from 204 ‰ to 219 ‰ between 1993 and 1999, which doubtless represents the major basic data for the PRSP.

A second and complementary set of health-related indicator includes measures of childhood nutritional status (weight for age or, more easily, weight for height). Although these are relatively easy to collect by means of surveys, they are more a reflection of the nutritional situation, thus economic, than the level of health in its broadest sense, and by the same token reflects the efficiency of health policy less well than the rate of child mortality, which remains more universally applicable.

In the education sector, the impact indicator used most often is the literacy rate of adults. However, this indicator remains both highly unreliable and slow to evolve as it results from the past rates of schooling and the possible schooling of adults, which is always limited. To get a better dynamic perspective on educational achievement, it may be better to focus on rates of schooling for children. They are obviously closer to measures of impact than are utilization rates in the health sector, since school is the necessary condition for access to knowledge, which is not true to the same extent for the utilization of public health care centres for a good level of health. Nevertheless, rates of schooling are not indicators which are directly representative of the acquisition of knowledge by children. Since they are contaminated by absenteeism and drop-out effects, more appropriate measures are primary-level exit or completion rates for children in any given age cohort. Even then, this does not adequately reflect the quality of the education received: for example, in Madagascar, while the schooling conditions for children were disrupted by the political upheaval, the success rate for the certificate of primary studies increased, no doubt due to a drop in the levels required. Generally, the increase in school attendance includes a risk of reducing quality if it is obtained at the price of an excessive increase in the number of pupils per teacher (either by double sessions, or by the increase of pupils in the classroom) or of the recruitment of under-qualified teaching staff: thus in Uganda, from 1991 to 1999, independent testing for knowledge in mathematics and English highlight a significant reduction of the level, as well as extreme dispersion from one school to the next.[13]

Monitoring the progress in education requires indicators which relate directly to the acquisition of knowledge by children who have received schooling. As hard as it is to establish knowledge testing and as much as these tests are subject to a cultural relativity when it is a question of secondary and higher levels of education, it nonetheless seems possible to have tests which are simple to administrate and interpret for knowledge acquired in primary education (reading, writing, basic arithmetic). Tests of this type do, indeed, exist (for example, through the Conference of Education Ministers in francophone countries). Moreover, as the Madagascar and Uganda reports indicate such tests are already effected in these countries.

Calculating impact indicators is not outside the domain of a strategy of collecting statistics based on a hierarchy of information.[14] However, an annual rhythm of impact indicator calculation, and thus evaluation, may be too frequent. Thus we recommend a two-phase procedure. Relying on impact indicators of necessity implies stretching out the evaluation interval upon which disbursement of variable tranche support is based, which means re-situating the budgetary support in a medium-term perspective. For example, three-yearly evaluation based on the evolution of the impact indicators would lead to the annual total of the variable tranche being determined for the subsequent three years.

Recourse to intermediate indicators will not disappear entirely. These are still the tools for monitoring state action, understanding the evolution of its impact and encouraging a modification of the policy. The regular production of a certain number of these indicators could even be a specific condition itself. The level achieved by the intermediate indicators would not act as a payment criterion, but establishing them on a regular basis could be such a criterion. Indeed, the interest of these indicators is to serve as a guide for state action and they

[13] In Ghana the high rates of completion appeared to go hand in hand with poor results in the knowledge tests (information supplied by F. Orivel).

[14] The cost of information necessary to establish impact indicators should not be overestimated. In a country like Burkina-Faso, according to the information collected by the National Statistics and Demography Institute, a DHS costs approximately twice the sum spent for only the realisation of the surveys relating to budgetary efficiency (like the prices of public markets) and opinion polls (for example among users of public education and health services and among economic operators on the State markets), which figure as a payment condition in the variable tranche and the interpretation of which has proven difficult.

would assume this role all the better if there were no financial interest to manipulate them. Hence, impact indicators would serve for the evaluation of the performance and for the modulation of aid whereas intermediate indicators would serve the monitoring of public action, which is the responsibility of the states. The European procedure would thus become more efficient with regard to the ownership of policies, without losing any information on their effective implementation.[15] [16]

4.3. From observation of indicators to the evaluation of results

The success of the reform of conditionality requires the credibility of the sanction. This implies that the evaluation of results is effected according to principles or rules known in advance. European Commission financing conventions provide a weighting of each indicator and a target that these indicators are supposed to achieve, as well as a scale of disbursement for the level of progress achieved relative to each target. This is typically from zero to 50 percent to 100 percent disbursement, although 100 percent disbursement may be achieved for outcomes that are close to but do not fully achieve the target. However, whatever the specification of the rules fixed in advance, a certain discretionary margin inevitably remains in the hands of the evaluator.

While it is essential that the weighting of the indicators be determined and known in advance, the actual weights accorded to education and health (or any other major component which is adopted) are ultimately a political choice which should logically result from the strategy of the fight against the poverty specific to each country. Within each of these two large domains, the problem of weighting is posed differently according to the nature and the number of indicators. If they are composite impact indicators, there are very few of them (sometimes

[15] As for the choice of intermediate indicators (outcome, output) which would be adopted for this public action monitoring (and which the state would undertake to produce), it could itself be guided not only by acquired experience but also by simple principles: relatively simple collection, low cost and reliability, and especially pertinence with regard to the final objective of improving health and education. It would therefore be justified to test the relevance of this structure by estimating the impact exercised by the variables corresponding to the effective acquisition of knowledge and the reduction in mortality (or morbidity).

[16] There is a problem of transition between the current system, with an annual evaluation of results triggering the payments, and the new system of a triennial evaluation founding future undertakings. We see that for the first three years of the new system, the aid will be defined then paid on the basis of the countries' commitment to achieve certain results at the end of the three years, the incentive to implement adequate policies being constituted by the risk of losing all future aid. The selection of the countries benefiting from the new procedure could be made on the basis of results obtained since 1999 within the context of the current procedure, which in some ways serves as a test .

only one per domain) and the problem of weighting can be resolved easily.[17] The question becomes thornier when there is a rather large number of indicators, which is generally the case with intermediate indicators aimed at outlining the different aspects of the policy implemented in the sector. The solution adopted has been, implicitly or explicitly, to give equal weight to each of them for reasons of simplicity, but with no logical foundation.

For health and education it would be desirable and possible to develop an econometric model testing the influence of the different factors reflected by the intermediate indicators on the final impact (reduction in child mortality or acquisition of knowledge). If such a model allows the identification of factors signifying the desired improvement, it also provides the structure for weighting indicators by means of the estimated coefficients. Though this type of information should certainly be used with care, it does, however, allow the public action priorities to be highlighted by indicating the most efficient channels for improving health or education.

There is a further problem of reference values. Indeed, to judge the improvement of whatever indicator (impact or intermediate); it must be compared to either an objective or target fixed in advance or a past value. In numerous cases, the new European procedure has resorted to targets defined in the hope that they facilitate evaluation by supplying an exact reference for the improvement of the policy. However, the use of targets revealed the following dilemma: either the government could set ambitious objectives to demonstrate the ambition of their policy and the results risk staying well below the objectives, or, on the contrary, they could aim to maximise their chance of success in relation to the targets by setting them at a relatively low level. The experience of Uganda contains, in this regard, many lessons: the review of the education sector by all the donors in April 2001 (a review carried out independent of the new European procedure) highlighted results well below the objectives, which led to the objectives being modified instead of aid being reduced.

A natural set of reference values for any country could, presumably, be derived from international comparison. Thus, “normal” values of each indicator would be established to which it would be suitable to come close or exceed. The international comparison of targets to achieve has much more sense for impact indicators than for intermediate indicators, which are

[17] For example in the domain of education, if we have an indicator for the acquisition of knowledge, it could simply be used to correct the completion rate for primary education.

often influenced by the individual characteristics of the countries. Moreover, it remains the case that by far the simplest approach would be to consider whether there is an improvement compared to the past value. However, the improvement potential of an indicator naturally depends on its initial level[18] For example, the potential for reducing child mortality is much stronger when the latter is very high than when it is very low, as the reduction tends to a limit close to zero. Similarly, the rate of vaccination coverage or the rate of schooling and acquisition of knowledge tend towards a limit equal to the unit (or 100 %): thus the closer we come, the more difficult it is to realise the same improvement.[19]

Exogenous factors and overall judgement

Finally, the evaluation of performances cannot not avoid taking into account exogenous factors which influenced the results. In an article echoing the work of Burnside and Dollar (2000), which established the new paradigm of the World Bank with regard to efficiency of international aid, Guillaumont and Chauvet (2001) showed that international aid contributes more to promoting growth when countries are recurrently subject to external shocks (particularly because the shocks render the implementation of economic policy more difficult and aid diminishes this effect). Consequently, it should be the “performance” of the country, corrected for the influence of exogenous shocks, as distinct from the gross results, which should serve as aid allocation criterion, since it both reflects the quality of the policy and takes into account the effect of exogenous factors.

In this perspective an overall judgement of performance in the area of health and education has to be formed in the light of the general political and economic context. In the present European conditionality the indexation of budgetary aid on several and distinct indicators appears to many national leaders and partners of development as a false objectivity. Conversely the subjectivity of a global judgement may be temperated by a joint evaluation of performance by the different donors.

---

[18] This has to be be evaluated by international comparison.

[19] There is also the problem of knowing if the past reference should be the last known year or an average of the two (three) last known years, or even a value calculated using the estimation of the past trend. If we exclude this last possibility due to the large number of years of statistical information required and the debates which the functional form of the estimation inevitably raises, the choice between two and three reference years, preferred to one, can be justified with a view to reducing the influence of past exogenous factors on the evaluation of performance.

Conclusion

The European Commission's new conditionality represents a bold and important contribution to the international debate on aid effectiveness. The underlying principles of this initiative attract a broad consensus of support: thus the necessity for an ownership of their economic policy by the recipient countries, a more efficient coordination of the different international donors and greater attention to the results of action undertaken, especially with regard to the reduction of poverty are also recognised. With a view to reinforcing the global efficiency of their development aid, the donors are today striving for a better allocation of their resources or a selectivity of their aid by taking into account the results obtained by the countries receiving their help. This is precisely what the European initiative allows. It is likely to increase the credibility of conditionality, and consequently of the sanction, due to its more progressive nature, which also improves the efficiency of the aid.

To succeed the new European Commission conditionality must be unambiguous. First, the initiative should concentrate exclusively on performance in the health and education sectors. Second, it must return to basics: the new conditionality must be based on real impact indicators and applied to future triennial aid commitments rather than to current payments. Third, intermediate indicators should be used for the annual monitoring of policy implementation, without their levels serving to modulate aid, and with only their production being a condition of annual payment; the justification for this control would be strengthened by a prior econometric analysis with a view to testing the impact of these intermediate variables on the final result. Fourth, the evaluation of impact should take into account exogenous factors and assume the responsibility for a global judgement of performance by including all the donors in the judgement in order to promote coordination. Fifth, continue to support the capacity for analysing the policies in the countries receiving aid in order to render the new procedure more efficient.

In this final perspective, a normal complement to the new conditionality would be long-term support for a strategy of priority information collection (in particular for DHS data generation and isometric testing for the acquisition of knowledge). This could occur under the auspices of institutional support (for example to the new national poverty observatories) and would allow the improvement in quality and regularity of the impact indicators. Thus established with a certain international guarantee, the impact indicators would serve more easily to guide

European support and, more broadly, the different sources of aid for the improvement of health and education.

References

Adam, C.S. and J.W. Gunning, “Redesigning the Aid Contract: Donors’ Use of Performance Indicators in Uganda”, *World Development*, vol. 30, 2002, pp. 2045-2056.

Adam C.S. and S.O’Connell, “Aid, Taxation and Development in Sub-Saharan Africa” *Economics and Politics,* vol. 11, 1999, pp. 225-254.

Burnside C. and D.Dollar, “Aid, Policy and Growth” *American Economic Review* , vol.90, no 4, 2000, pp. 847-868.

CERDI, *Etude sur l’introduction d’indicateurs de résultat en matière d’appui aux programmes d’ajustement structurel dans les pays ACP*, réalisée à la demande de the European Commission, Clermont-Ferrand, 2002.

Collier P., Guillaumont P., Guillaumont Jeanneney S. et J.W. Gunning, "Redesigning Conditionality", *World Development*, vol. 25, 1997, pp. 1399-1407.

Collier P. and J.W. Gunning “The IMF’s Role in Structural Adjustment”, *Economic Journal,* vol.109, 1999, pp.F634-651.

Commission of the European Communities , Communication from the Commision to the Council and the European Parliament, *Community Support for Economic Reform Programmes and Structural Adjustment: Review and prospects, 2000, Brussels, COM(2000)yyy final.*

Guillaumont P. and L. Chauvet, “ Aid and Performance: A Reassessment “, *Journal of Development Studies,* vol. 37, 2001, pp. 66-92.

Guillaumont P. and S. Guillaumont Jeanneney (sous la direction de), *Adjustment and development. The experience of the ACP countries*, Paris, Economica, 1994.

Guillaumont P and S. Guillaumont Jeanneney "La conditionalityà l'épreuve des faits", in R. Rainelli (ed.), *La négociation commerciale et financière internationale,* Paris, Economica, 1995, chap. 8, pp. 185-208.

Guillaumont P and S. Guillaumont Jeanneney *Propositions pour un nouveau type de conditionnalité*, Commission Européenne, Direction Générale du Développement, Division Programmation et soutien à l'ajustement structurel, 1995a, octobre.

Gunning J.W., "Rethinking Aid", in B. Pleskovic and N. Stern, *Annual World Bank Conference on Developmenr Economics,,* Washington, World Bank, 2000, pp.125-144.

IMF, *The Modalities of Conditionality – Further Considerations,* prepared by the Policy Development and Review Department, January 8, 2002, Khan M.S. and S. Sharma, “IMF Conditionality and Country Ownetrship of Programs” IMF *Working Paper* 01/142, September 2001.

Lerrick A. et A. Meltzer, “Grants: A Better Way to Deliver Aid”, Carnegie Mellon, Gailliot Center for Public Policy, *Quarterly International Report,* January 2002.

Meltzer A. and others, *International Financial Institutions Advisory Commission Report,* Washington, DC, 2000.

Mussa M. et M. Savastano, “The IMF Approach to Economic Stabilization”, IMF *Working Paper* 99/104, 1999.

Tirole J. “Financial Crisis, Liquidity Provision and the International Monetary System” Paolo Baffi Lecture on Money and Finance, Bank of Italy, 2001.

World Bank, *Assessing Aid: What Works, What Doesn’t, and Why,* A World Bank Policy Research Report, Oxford University Press, 1998.

Appendix

Table 1- Classification of programs 1999-2001

Traditional conditionality New Conditionality	Maintain	Abandon
Definition of indicators	CFA** *Gabon*[20] *Côte d'Ivoire*[21]* *Gambia*[22]*	Kenya** Papua New Guinea** *Ghana** *Jamaica (SERP II)**
Use of indicators	Cameroon** Guinea Bissau** Lesotho** Mozambique** Niger** Rwanda** Sao Tome and Principe** Tanzania** Zambia** *Ethiopia*[23]* *Republic of Guinea*[24]*	Benin *** Benin** Burkina Faso *** Burkina Faso** Djibouti* Gabon* Madagascar *** Mali *** *Mali*[25]* *Mauritania*[26]* Uganda** Sierra Leone** Chad**

* Financing proposals 2001 (10)
** Financing proposals 2000 (19)
*** Financing proposals 1999 (3)
in italics: *programmes 2001*

---

[20] Variable tranche determined by progress achieved in satisfying certain budgetary performance criteria and the availability of indicators of results in the social sectors.
[21] One tranche paid upon approval by the IMF of new FRPC programme and one tranche linked to meeting conditions of this programme and arriving at an agreement on performance indicators for the social sectors and public administration.
[22] Payment in one tranche in accordance with the objectives of the PSRP.
[23] Payment according to results in public administration, social sectors, and private sector enabling environment.
[24] Fixed tranche according to World Bank and IMF evaluation plus a variable tranche based on performance in public administration and an agreement on the indicators of health, education and road maintenance.
[25] Payment linked to perfromance indicators in public administration, health, education and governance.
[26] Payment of variable tranche linked to perfromance in public administration and the social services.

Table 2 – Comparable progress of the programs

	Program number	Years and designation of the program	Status of the evaluation
Benin	2	PAS 3 1999-2000	In progress during mission (complete)
		PARE 2001	
Burkina-Faso	2	PAAS 1999-2000	Evaluated
		ABRP 2001	Evaluation in progress during mission (complete)
		New program in preparation	
Madagascar	1	PAS 2 Initially 1998-1999 (signed in October 1999) and then designated PAS 2 1999-2000	Evaluated
Uganda	1	PABS 4 signed in August 2000	Still not evaluated at the time of the mission (payment of the variable tranche envisaged for November 2002

PAS and PAAS : Programme d'appui à l'ajustement structurel, PARE : Programme d'appui aux réformes économiques, ABRP : Appui budgétaire aux réformes économiques. PABS : Poverty alleviation budget support

Table 3 - Nature of the indicators of results
for European budgetary aid in four countries

	Number of indicators				Nature of the conditions			Weighting of each sector			Weighting within each sector	Global success threshold (bonus)	Taux de déboursement après évaluation
	Total	of which			Definition, collection and analysis of indicators	Favourable change	Target achieved	GB	Ed.	Health			
		GB	Ed.	Health									
Benin													
PAS III 1999-2000	9	-	-	9	4	5		-	-	1	no ex-ante equals ex-post	no	67 %
PARE 2001	13	3	5	5	3	1	10	1/3	1/3	1/3	equal	no	-
Burkina Faso													
PAS 1999-2000	15	5	3	7	4		11	0.5	0.15	0.35	equal	no	79 %
ABRP 2001	13	3	5	5	3	1	9	0.5	0.25	0.25	equal	no	60 %
Madagascar													
PAS II	54*	25	19	10	54**			1/3	1/3	1/3		80 %	71,4 %
Uganda													
sPABS4	(12)	(3)	5	4							equal	80 %	-

* Of which 15 "neutralised" and one indicator of the % of road funds allocated to rural tracks.
An indicator is said to be "neutralised" when it has not been taken into account to determine the sum of the variable tranche.
GB = budgetary management; Ed = education.