

*METHODOLOGY FOR EVALUATIONS OF BUDGET SUPPORT
OPERATIONS AT COUNTRY LEVEL*

Issue Paper

May 2008

Assignment for the European Commission





Development
Researchers'
Network
Italy



Aide à la Décision Economique
Belgium



ECO-Consulting Group
Germany



Nordic Consulting Group
Denmark



ECORYS
The Netherlands

A consortium of
DRN, ADE, ECO, NCG, ECORYS
c/o DRN, leading company

Headquarters

Via Ippolito Nievo 62
00153 Rome, Italy
Tel +39-06-581-6074
Fax +39-06-581-6390
mail@drn-network.com

Belgium office

Square E. Plasky 92
1030 Brussels, Belgium
Tel: +32-2-7324607/7361663
Fax +32-2-706-5442
bruxelles@drn-network.com

This assignment was commissioned by:

the Evaluation Unit common to:

EuropeAid Co-operation Office,
Directorate General for Development and
External Relations Directorate-General

This assignment was carried out by: Enzo Caputo of Development Researchers' Network – DRN (team leader); Andrew Lawson of the Overseas Development Institute - ODI; and Martin van der Linde of ECORYS. The team was assisted by Diego Ruiz (DRN).

Camilla Valmarana was the Evaluation Contract Manager.

The assignment was managed by the evaluation unit who also chaired the reference group composed by members of the services (AIDCO, RELEX, DG DEV, DG ECFIN, DG BUDG).

The opinions expressed in this document represent the authors' points of view which are not necessarily shared by the European Commission or by the authorities of the countries concerned.

PRELIMINARY REMARKS

This Issue Paper presents and analyses in detail the methodological issues underlying the proposed approach to the evaluation of budget support operations.

It addresses issues of both a conceptual and methodological nature and provides the necessary background to justify and explain - from a theoretical and methodological point of view – the choices made and approach proposed.

In the authors view, the paper is therefore not intended as a prescriptive / normative document. On the contrary it is presented – and should be viewed - as an instrument for debate and discussion to explain the current use of the theory-based approach and impact evaluation as conceived in the international discussion forums. This is why the paper is also relatively open to possible integrations and complementarities with other approaches to policy evaluation and evaluation of GBS/SBS.

The paper is structured in two sections: the first section provides a literature review of approaches to budget support evaluation and a discussion of the main related methodological issues; the second section presents the origin and justification for the proposed intervention logic and methodological approach developed for the evaluation of budget support operations, namely the three-step approach. The methodological section is completed by the description and analysis of different approaches to the construction of counterfactuals and concludes by recommending the combined use of qualitative and quantitative methods.

This Paper has been drafted by a team including the following: Enzo Caputo of Development Researchers' Network – DRN (team leader); Andrew Lawson of the Overseas Development Institute - ODI; and Martin van der Linde of ECORYS, assisted by Camilla Valmarana and Diego Ruiz (DRN), and Esther van der Meer (ECORYS). They are indebted to all the members of the Reference Group for the comments and guidance they provided. However, final responsibility for the content of this document rests with the team.

TABLE OF CONTENTS

SECTION ONE - LITERATURE REVIEW

1	WHAT IS THE OBJECT OF A BUDGET SUPPORT EVALUATION?.....	1
1.1	Definitions of General and Sector Budget Support	1
1.2	The evolution of general and sector budget support in the last decade	2
1.3	Distinguishing GBS and SBS	2
2	APPROACHES AND METHODOLOGIES IN THE RECENT GBS/SBS EVALUATIONS UNTIL 2007	4
2.1	Two approaches to GBS/SBS evaluations	4
2.2	GBS/SBS reviews which Do Not Use an Explicit Intervention Logic	4
2.3	GBS/SBS Evaluations which Use an Explicit Intervention Logic	6
2.3.1	<i>The Evaluation Framework of General Budget Support (EF)</i>	7
2.3.2	<i>The Enhanced Evaluation Framework (EEF)</i>	8
2.3.3	<i>Other recent Joint Evaluations</i>	9

SECTION TWO - THE INTERVENTION LOGIC AND OTHER KEY ISSUES

3	THE RATIONALE FOR BUILDING AN INTERVENTION LOGIC	10
3.1	The basics of an intervention logic	10
3.2	A Closer Integration between GBS/SBS Evaluation and assessment of the Government's strategy	11
3.3	The Comprehensive Evaluation Framework	13
4	THE THREE STEP APPROACH TO THE APPLICATION OF THE FRAMEWORK	15
4.1	An approach to the evaluation that fits the intervention logic	15
4.2	Key issues for a three-step approach	15
4.3	Overview of the three-step approach	16
4.3.1	<i>Key components and overlapping</i>	16
4.3.2	<i>Peculiarity of Step Three</i>	18
4.4	The Evaluation Questions	19
4.5	Adjusting the focus of the evaluation to the country's level of aid dependency	21
4.6	Adjusting the focus of the evaluation to the type of budget support	23
5	COUNTERFACTUALS AND OTHER EVALUATION TOOLS.....	24
5.1	Counterfactuals	24
5.1.1	<i>Counterfactual analysis and impact evaluation</i>	24
5.1.2	<i>Different approaches to counterfactuals</i>	24
5.2	Quantitative and qualitative methods	27
5.2.1	<i>Mixed methods</i>	27
5.2.2	<i>Statistical analyses</i>	28
	ANNEX 1: BIBLIOGRAPHY	29
	ANNEX 2: SUMMARY OF THE MAIN RECENT DEFINITIONS OF BUDGET SUPPORT	32
	ANNEX 3: USE OF COUNTERFACTUALS IN THE EVALUATION OF COMPLEX PROGRAMMES	34

ACRONYMS

CB	Capacity Building
CEF	Comprehensive Evaluation Framework
DAC	Directorate for Aid Coordination (OECD)
DFID	Department for International Development (UK)
EC	European Commission
ECFIN	EC Directorate General for Economic and Financial Affairs
EEF	Enhanced Evaluation Framework
EF	Evaluation Framework
EQ	Evaluation Question
G/S	General and/or Sectoral
GBS	General Budget Support
GS	Government Strategy
IDD	International Development Dep't (Un. Birmingham; has led the JEGBS)
IL	Intervention Logic
JC	Judgement Criteria
JEGBS	Joint Evaluation of General Budget Support
MDB	Multilateral Development Banks
MDBS	Multi-Donor Budget Support
MTEF	Medium-Term Expenditure Framework
ODI	Overseas Development Institute
OECD	Organisation for Economic Cooperation and Development
PBL	Policy-Based Lending
PFM	Public Financial Management
PRS	Poverty Reduction Strategy
PRSC	Poverty Reduction Support Credit
PRSP	Poverty Reduction Strategy Paper
SBS	Sectoral Budget Support
SIDA	Swedish International Development Agency
SPA	Strategic Partnership with Africa
SWAp	Sector Wide Approach
SWAP	Sector Wide Approach Programme (government led)
TA	Technical Assistance

SECTION ONE - LITERATURE REVIEW

1 WHAT IS THE OBJECT OF A BUDGET SUPPORT EVALUATION?

Before addressing the methodology for evaluation of budget support, it is first necessary to identify a definition of budget support that fits the scope of the ToR and is compatible with the terminology and interpretation currently adopted by the European Commission (EC) and accepted by the OECD and the main International Aid Agencies.

1.1 DEFINITIONS OF GENERAL AND SECTOR BUDGET SUPPORT

According to the definitions given by the EC and the OECD/DAC, the essential characteristic of budget support – whether general or sectoral - is that funds are channelled directly to the partner government's general treasury account and are subsequently deployed according to the country's own allocation, procurement and accounting systems (see box 1).

Box 1: OECD – DAC and EC Definitions of Budget Support

Institution and Source	Definition
<p>EC AIDCO, DEV, RELEX <i>Aid Delivery Methods: Guidelines on the Programming, Design & Management of General Budget Support</i>, January 2007</p>	<p>Budget support is the transfer of financial resources of an external financing agency to the National Treasury of a partner country, following respect by the latter of agreed conditions for payment. The financial resources thus received are part of the global resources of the partner country, and consequently used in accordance with the public financial management system of the partner country.</p> <p>Budget support as described above may be provided as either:</p> <ul style="list-style-type: none"> - <u>General Budget Support</u>, representing a transfer to the national treasury in support of a national development or reform policy and strategy; or - <u>Sector Budget Support</u>, representing a transfer to the national treasury in support of a sector programme policy and strategy.
<p>OECD/DAC <i>Harmonizing Donor Practice for Effective Aid Delivery: Volume 2; Budget Support, Sector-Wide Approaches and Capacity Development in Public Financial Management</i>, 2006</p>	<p>Budget support is defined as a method of financing a partner country's budget through a transfer of resources from an external financing agency to the partner government's national treasury. The funds thus transferred are managed in accordance with the recipient's budgetary procedures. [...] In the case of <u>general budget support</u>, the dialogue between donors and partner governments focuses on overall policy and budget priorities, whereas for <u>sector budget support</u> the focus is on sector-specific concerns.</p>

A more detailed list of main recent definitions of budget support is provided in Annex 2.

Within this general definition of budget support, a distinction can be made between General Budget Support (GBS) and Sector Budget Support (SBS):

- **General Budget Support** is defined as the form of budget support of which the purpose is to contribute to the implementation of comprehensive goals set out in the national development or poverty reduction strategy. The policy dialogue is focused on overall policy objectives and overall budget priorities. Funds provided through General Budget Support are by definition un-targeted¹.

¹ Targeting (term adopted by the EC) – or earmarking (term adopted by various other development partners) - is generally intended to denote the consequence of an agreement between government and GBS/SBS

- **Sector Budget Support** is defined as a form of budget support of which the purpose is to contribute to accelerated progress towards the government's goals within a specific sector. The policy dialogue is focused on sector-specific issues and concerns. As in GBS, funds are generally un-targeted. There may however be agreements related to the level of public spending in the sector concerned and sometimes agreements on the composition of spending in the sector.

The above-mentioned definitions refer to the financial flows and the focus of the policy dialogue. On top of that, budget support is usually also characterised by efforts to strengthen harmonisation between development partners (DPs), to enhance the alignment of external support with the country policies and systems, and to reduce the transaction costs of providing and receiving external assistance. In a number of cases, budget support is also accompanied by provision of capacity development assistance.

Following on from the definitions mentioned above, the evaluation of General or Sector Budget Support *will relate to a comprehensive form of development assistance of which the **objective** is to support general or sectoral government strategies, which includes a number of **inputs** (funds transferred directly to the partner government's Treasury account; policy dialogue and conditionality; and often capacity development assistance), and an advanced **partnership framework** (DP alignment with government system and harmonisation).*

1.2 THE EVOLUTION OF GENERAL AND SECTOR BUDGET SUPPORT IN THE LAST DECADE

The above definitions refer to the specific forms of GBS and SBS which took shape in the late 1990s, as offsprings of the HIPC debt reduction initiative and the funding of sector-wide approaches. The difference between these new forms of Budget Support and the previous versions is related to their (implicit) purpose as well as to the nature of the relationship between the DPs and recipient governments. In the 'old' approach the relationship was rather narrowly contractual (funds *vis-à-vis* macro stabilisation and liberalisation); in the 'new' approach the relationship relies on a broader partnership entailing acceptance of the respective roles of the partners and sharing of responsibilities².

Although the change in the purpose and design of budget support is important, it should be stressed that the distinction between "old" and "new" budget support is not always clear-cut. Evaluators should therefore analyse all GBS or SBS programmes provided within the specified evaluation time period, with a view to assessing their combined effects. If, on closer inspection, the budget support operations are found to be less partnership-based than had been thought (e.g. more akin to "old" budget support operations), then this is likely to be an important explanatory factor in determining the outputs, outcomes and impacts of budget support but it should not be a reason for excluding such operations from the evaluation.

1.3 DISTINGUISHING GBS AND SBS

A comprehensive analysis and discussion of the differences and links between GBS and SBS was provided by the Strategic Partnership with Africa (SPA) workshop on sector budget support held

providers regarding the actual destination of the funds transferred. Such an agreement may be more or less binding for the government, but it should never be confused with the focus of the dialogue aimed at strengthening poverty reduction and sectoral development objectives and targets.

² This new approach to Budget Support is also referred to as Partnership Budget Support, to stress the new partnership framework built since the second half of the '90s, including the post-Washington consensus, the HIPC and the Poverty Reduction Strategy Papers.

in Dublin in October 2005. The outcome of that workshop suggested that there should be no pronounced conceptual separation or divergence between these two forms of budget support as developed in recent years.

The workshop concluded that *“the sharp distinction [...] between GBS and SBS is not reflected in donor practice. It is more realistic to describe budget support as a spectrum. At one extreme is GBS with dialogue and conditions focused only on macro and cross-sector issues. At the other extreme is SBS focused only on sector-specific issues. In between is GBS with sector conditions and dialogue and those SBS operations which include some macro and cross-cutting conditions and dialogue”*³.

While there is a difference in the policy focus, the two forms of budget support are consistent with each other and use the same partnership approach based on alignment and harmonisation.

In keeping with these realities, the evaluation framework and approach proposed in section 2 are generic, and may be applied to three eventualities:

- situations where only General Budget Support is provided;
- situations where General Budget Support is provided alongside one or more Sector Budget Support arrangements;
- situations where Sector Budget Support is provided on its own.

The scope of the evaluation should be determined by the particular circumstances of the country and the time period to be evaluated. Where GBS and SBS are provided together, it is worthwhile to evaluate them at the same time because they contribute jointly to the changes in the funding composition of the budget, in the overall framework for policy dialogue, in the institutional context for public spending and in the overall partnership framework. The GBS evaluation will then focus on the aggregate level and the SBS evaluation at sector level, where the specific effects of the SBS arrangements will be identified. At the aggregate level evaluators should not attempt to separate the precise effects of GBS and of SBS if these are provided simultaneously. Although it is recognised that different programmes might be due for evaluation at different times, it is proposed that – as far as possible – budget support operations be evaluated jointly⁴.

In the case of a ‘stand-alone’ SBS evaluation, the Evaluation Questions have to be applied at both aggregate and sector levels. The presence of two or more SBS arrangements provides a ready source of comparative experiences from which valuable evaluation lessons can be drawn.

³ SPA - Strategic Partnership with Africa. (2005), Sector Budget Support: A Note from the Dublin Workshop of SPA Working Groups, 5-6 October 2005.

⁴ The different stages of the various BS operations should be clearly recognised and stated, and it is likely that these different stages will be reflected in the findings at the different levels of the framework. The actual choice in terms of timing will have to be made on a country-by-country basis.

2 APPROACHES AND METHODOLOGIES IN THE RECENT GBS/SBS EVALUATIONS UNTIL 2007

2.1 TWO APPROACHES TO GBS/SBS EVALUATIONS

Recent and on-going evaluation work to assess the effects of budget support might be broadly divided into two main categories depending on the use - or lack - of an explicit GBS/SBS intervention logic as the main evaluation tool.

Studies that do not employ an explicit intervention logic include reviews and assessments focusing on the degree of attainment of a number of objectives, regardless of the possible causality chains or without spelling out the causal linkages postulated by the implicit thinking behind GBS/SBS programmes in recent years. Among them are papers that focus either on the assessment of specific cases, or on the overall review of a number of cases to draw general lessons and guidelines.

Studies that do employ an explicit intervention logic include reviews and assessments focusing not only on the progress made towards the attainment of given objectives but also on the underlying processes (and causality chains) that have led to that progress.

2.2 GBS/SBS REVIEWS WHICH DO NOT USE AN EXPLICIT INTERVENTION LOGIC

The following box presents a brief overview of the first group.

Box 2: GBS/SBS Reviews that Do Not Use an Explicit Intervention Logic⁵

OECD “Harmonizing Donor Practices for Effective Aid Delivery, Volume 2: Budget Support, Sector-wide Approaches and Capacity Development in Public Financial Management”. This report relies on a broad experience assessment to identify guiding principles and good practices in providing general budget support and support to SWAs. In doing so, it acknowledges: i) the need to strengthen partnerships through increased collaboration not only between donors and partners but also between the various donors; and ii) the special relevance of public financial management issues for both of these modalities of aid delivery. The report puts forward a set of good practices on how donors can best deliver budget support in ways that maximise its developmental benefits while reinforcing partners’ capacity to achieve aggregate fiscal discipline, strategic allocation of funds, value for money, and probity in the use of public money – these are all key objectives of public financial management systems. It also provides donors with benchmarks for behaviour when providing budget support. Good Budget Support practices are grouped under four main pillars: i) supporting ownership (refrain from targeting support, reflect partner country priorities, focus on results); ii) enhancing public financial management (PFM) performance and accountability (follow good practices in PFM diagnostic and assessment work, directly support the capacity development of partner PFM systems, avoid undermining country systems); iii) reducing transaction costs (streamline conditionality, rationalise fiduciary assessments, align processes, tap the potential of joint donor frameworks, time disbursements to facilitate the smooth execution of budgetary payments); iv) enhancing predictability and reducing volatility (programme budget support over several years, align support with partner country budget cycles, design conditionality to enhance the predictability of disbursements, time disbursements in a predictable manner, avoid stop-and-go cycles and allow for graduated responses, build public support).

⁵ This box builds and expands on the ‘Recent and ongoing Evaluations of Budget Support’ box presented in S. Koeberle S., Stavreski Z. (2005) *Budget Support: Concept and Issues*, The World Bank.

World Bank (2006). “A Stocktaking of Poverty Reduction Support Credits”, J. Factora in *Budget Support as More Effective Aid? Recent Experiences and Emerging Lessons*, S. Koeberle, Z. Stavreski, J. Walliser (eds). This paper outlines the main findings of the review of the WB’s experience with Poverty Reduction Support Credits (PRSCs)⁶ implementation and measures the extent to which the PRSCs have made progress in achieving their objectives and expectations. The paper focuses on the analysis of progress made towards the PRSCs overarching goal of supporting the PRSP approach and of reinforcing country ownership through improvements in three main areas that correspond to the mutually reinforcing objectives of the PRSCs: i) operationalisation and implementation of the PRSPs; ii) improvement of resource predictability; iii) aid harmonisation. With very few countries having completed their first series of PRSC operations, the general thrust of the paper is on design and implementation issues. Although the paper does not attempt to assess development impact, it highlights preliminary medium-term outcomes from the more mature PRSC programmes through the presentation of lessons learned, key issues, emerging practices and recommendations.

Strategic Partnership for Africa – Budget Support Working Group (2005). *Survey of the Alignment of Budget Support and Balance of Payments Support with National PRS Processes*. The report presents the main findings from a monitoring exercise undertaken in 2004 to stimulate dialogue between donors and governments on the poverty reduction strategy (PRS)-alignment issue, and to help in identifying good and bad practices that could be taken into consideration in ongoing discussions at country level. The primary objective of the report is to describe the current status of efforts to align the delivery of budget support and balance-of-payments support with national policy-making processes, budget cycles and systems for reviewing progress. Secondary objectives are to promote policy dialogue on this subject, disseminate good practices and identify countries where follow-up activities by the Strategic Partnership for Africa (SPA) might bring benefits. The findings are presented in three sections, corresponding to the three parts of the questionnaire completed respectively by 1) donor and government representatives jointly; 2) individual donor representatives on their own; and 3) government representatives on their own. The first section gives basic data on countries’ PRS and related review processes, and an overview of alignment issues based on the joint responses. The second section describes the programmes of support, and goes into greater detail on a range of alignment issues, drawing on the questionnaire returns and documents provided by the donor representatives. The third section is based on the assessments of progress in different dimensions of alignment and harmonisation achieved by the government representatives.

USAID (2005) “General Budget Support: Key findings of Five USAID studies”, PPC Evaluation Paper No.7. This paper synthesises the findings of five evaluations in countries (Malawi, Mozambique, Nicaragua, Tanzania and Timor-Leste) in which GBS had been used or was being considered. The studies are part of the Evaluation Office’s field efforts to analyze conditions needed for successful general budget support and to capture the effectiveness of General Budget Support as an aid instrument. More specifically the evaluations aimed at exploring how and when using GBS as an assistance mechanism would be appropriate, and also at identifying conditions that contribute to either success or failure. GBS was assessed in two ways: in one case the analysis focused around the question “what minimum country conditions are necessary before donors can consider GBS?”; in the second case the main question to be answered was “has GBS led to the successes claimed for it and are country capabilities improving?” It is important to note that the studies do not evaluate the impact of GBS on poverty alleviation. Instead they focus on perceived benefits within specific conditions and timeframes. Findings are grouped into six main issues: i) host-country ownership of development programmes; ii) budget allocation, performance assessment, and disbursement predictability; iii) donor coordination and harmonisation; iv) transaction costs; v) management and technical capacity; vi) fiduciary risk.

⁶ In the WB context, budget support is characterised by the Poverty Reduction Support Credits – a programmatic approach to development policy lending that typically consists of three or four annual, single-tranche operations, aimed at supporting the government’s PRS medium-term development objectives.

Most of the studies mentioned above focus on two sets of results:

- results directly dependent on aid activities and referring mainly to changes in donor behaviour such as alignment and harmonisation, or to absolute or relative increases of aid in the national budget, disbursement predictability, and so on.
- results which might have a strong or weak causal relationship with donor behaviour but are certainly not directly dependent on it. These include improved budget management and better links between policy and budget; or, to go further, they also address issues related to service delivery and development impacts.

The approach to the assessment of the two groups of results is logically different. If for the first set of results one may focus on a before-and-after analysis of aid flows and policies, for the second set of results a more complex approach is needed: first an assessment of what changes occurred and what were the determining factors; and second an accurate analysis aimed at identifying the contribution of the aid-related inputs to such factors. It must be said that the studies mentioned above do not go very far in such analyses as they concentrate on the assessment of the first group of results.

2.3 GBS/SBS EVALUATIONS WHICH USE AN EXPLICIT INTERVENTION LOGIC

As already mentioned, the reviews and assessments discussed in this section focus not only on progress towards attainment of intended objectives but also on the role and weight of the GBS/SBS inputs and on the mechanisms through which they contributed to that progress. To do this they build an intervention logic - or programme theory - as a framework for unpacking the expected GBS/SBS process and therefore allow an analysis of the extent to which, and the circumstances under which, Budget Support has been relevant to the context, and efficient and effective in achieving sustainable impacts on poverty reduction and growth; and what are the expected intermediate changes (according to the theory) in those government systems regarded as crucial mechanisms through which poverty outcomes and impacts will be improved.

Papers reviewed include:

- Lawson, A. and Booth, D. (2004). *Evaluation Framework for General Budget Support* commissioned on behalf of the OECD-DAC Evaluation Network. This report presents an Evaluation Framework (EF) intended to guide the conduct of joint evaluation work on General Budget Support (GBS) at country level. It is intended as a practical tool that can be used to assess whether GBS is a relevant, efficient, effective and sustainable mechanism for poverty reduction. The Framework is based on a flow of diagram of the Logical Framework type, which spells out the causal linkages posited by the theory behind recent GBS programmes;
- Nilsson, M. (2004). *Effects of Budget Support – A Discussion of Early Evidence*. UTV Working Paper 2004:4. Stockholm: SIDA. This paper summarises existing findings, lessons and recommendations in relation to the different levels of the above-mentioned evaluation framework and provides donors with early evidence and lessons on the effectiveness of GBS as an aid package (not just a financing modality);
- IDD and Associates (2007) *Joint Evaluation of General Budget Support 1994–2004*, commissioned by a consortium of donor agencies and seven partner Governments under the auspices of the DAC Network on Development Evaluation⁷. This work undertakes a comprehensive review of the

⁷ Documents and reports reviewed include the technical proposal submitted in response to the invitation to tender (2004); the Inception Report (2005); the Synthesis Report (2006) and the Note on Approach and Methods (2007).

previous methodology and sets up an Enhanced Evaluation Framework (EEF) which builds on the Evaluation Framework (EF) developed by Lawson and Booth and tries to integrate and develop it.

- ODI and Ghana Centre for Democratic Development (2007). “Joint Evaluation of Multi-Donor Budget Support to Ghana”. The report assesses the impact of GBS in Ghana building on the EF-EEF experience, and analyses not only what has happened but also how, why, and in what context the Ghana Multi-Donor Budget Support (MDBS) has attained its objectives.

2.3.1 The Evaluation Framework of General Budget Support (EF)

The EF is the starting point of the methodological approach under review in this section. Developed in 2004 by Lawson and Booth for the OECD-DAC, it was intended as a practical tool to be used as a guide for country-level joint evaluations. To this end, the expectations, priorities and objectives of budget support were unpacked and discussed in the EF which represented an “effort to set out in a systematic way the principal claims made on behalf of General Budget Support as a modality of poverty-oriented aid, spelling out the implied causal linkages in Logical-Framework fashion”⁸.

To develop the EF the authors analysed the rationale for GBS, its objectives and the set of hypotheses underlying the process through which it is postulated that GBS might achieve its objectives. To this end the proposed methodology focuses on a qualitative analysis of the public-expenditure processes, which is supposed to be influenced by GBS through both its institutional and flow-of-funds effects. Key features of the EF include:

- ✓ The standard five-level logical sequence to establish the cause-and-effect links and the time dimension of the effects:
 - Level 1: **inputs** by GBS donors, including funds, policy dialogue, conditionality, technical assistance / capacity building (TA/CB), alignment to government policies and systems, harmonisation between donors;
 - Level 2: **immediate effects** of the inputs on the relationship between aid, the national budget and national policy processes;
 - Level 3: **outputs**, that is consequent changes in the financing and institutional framework for public spending and public policy;
 - Level 4: **outcomes**, that is interactions between the public sector and the wider economy and society, specifically with regard to the proximate determinants of poverty reduction;
 - Level 5: **impacts**, in terms of empowerment of the poor and improvement of their real incomes.
- ✓ Provision of detailed guidelines for research questions and approaches at each level of the framework, based on assessing whether the postulated effects of GBS are present and asking additional questions relating to attribution and the counterfactual.
- ✓ Reliance on a pragmatic combination of quantitative and qualitative approaches, with cross-checking and triangulation.
- ✓ Treatment of factors outside the main hypothesised chain of effects as assumptions and risks (although these are to be explicitly considered in asking “why/why not?” questions related to attribution).

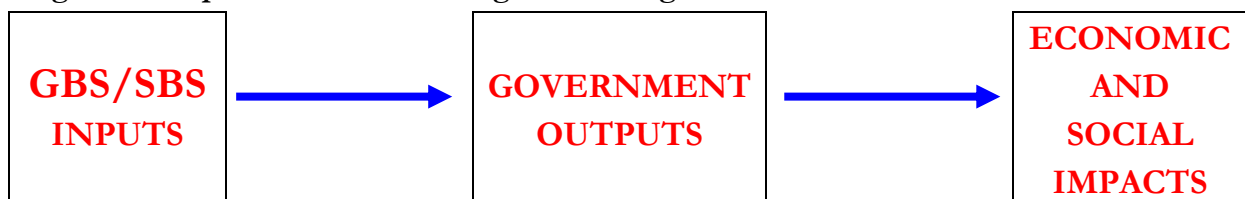
⁸ Lawson, A., Booth, D. “Evaluation Framework for General Budget Support”, ODI, 2004

2.3.2 *The Enhanced Evaluation Framework (EEF)*

The work undertaken by IDD and Associates in 2006-7 in the framework of the Joint Evaluation of General Budget Support led to the further development of the Evaluation Framework (EF). The thinking and proposed improvements in relation to the EF have been grouped into different categories as shown below.

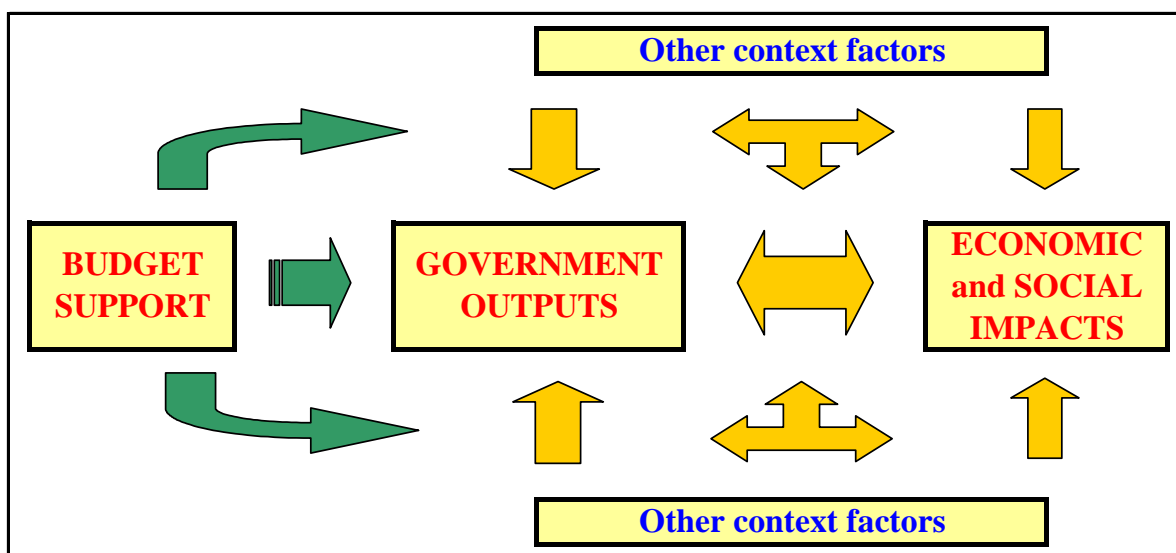
A first category addresses the theory underlining the respective approaches. The EF highlights the linear relationship between GBS/SBS inputs, expected government outputs and desired socio-economic impacts. Figure 1 shows a simplified scheme illustrating an intervention logic according to a standard linear model.

Figure 1: Simplified intervention logic according to a linear model



According to Figure 1 GBS/SBS inputs have a causal effect on government outputs, and the latter have a causal relationship with economic and social outcomes and impacts. Indeed the relationships are much more complex than that. Government outputs are determined by a complex range of factors, some of which (not necessarily those that are most significant for change, and not the same everywhere) may be influenced by GBS/SBS inputs. In turn, economic and social impacts are determined by an even more complex range of factors, of which some may be related to the government outputs influenced by GBS/SBS inputs. Figure 2 represents, in a simplified form, this complex relationship.

Figure 2: The representation of a more complex relationship between GBS/SBS inputs and the various factors affecting the change.



The EF tries to take into account the various factors affecting the change, including a detailed list of assumptions. The assumptions as such, however, can only be addressed as conditions or factors either facilitating or hampering the linear relationship from GBS/SBS inputs to outputs and outcomes. They are not analysed *per se* as possible determinants of or obstacles to the change.

The JEGBS team has partly identified such limits of the EF and has proposed putting in place a broader policy analysis approach based on five categories of analysis (PFM, Partnership, Institutional, Macro-economic and Poverty). These five analyses should facilitate understanding of the actual drivers of change as a basis for identifying the contribution of GBS/SBS. This approach gave some results, particularly in some countries, but was hampered in general by a weak methodological link between the two levels of assessment (the five analyses and the GBS/SBS intervention logic), and also by the structure of the report which emphasised the search for predetermined causality chains along the programme framework.

Other significant attempts by the JEGBS team to enhance the EF have drawn on certain analytical tools and a wider understanding of the areas of influence of the GBS/SBS.

On the analytical side the most significant innovation is related to the attempt to identify some particularly sensitive causality chains across the logical model, which need to be analysed in depth to assess their significance, draw lessons on the way in which they operate, and develop modalities for improving their effectiveness. In practice the attempt to identify a standard causality map to guide the analytical work of the evaluators might have led to a strengthening of the limits of the linear approach, at the cost of jeopardising the policy analysis approach. Much good analytical work has been done to test such predefined links, but given the wide degree of variation between the local situations and contributory factors, no strong conclusion was reached nor were any general lessons identified. Probably there has been pressure to analyse the relative importance of single inputs rather than the soundness of the whole system.

On the areas covered by the model, the JEGBS team has stressed the need to give stronger emphasis to the policy process and actual policy changes in the EEF. Indeed, GBS/SBS in recent years, compared to the previous structural adjustment experience, has focused much more on the linkages between policies and public expenditure. It gives equal importance to both public finance and policy management, and particularly to the government system that ensures their close linkage, including central government and line ministries, budget allocation procedures, policy implementation, and results-based monitoring.

2.3.3 Other recent Joint Evaluations

The Joint Evaluation of Multi-Donor Budget Support to Ghana is the latest evaluation that builds on the methodology developed through the Evaluation Framework (EF) and, in part, the Enhanced Evaluation Framework (EEF). Some important insights from the evaluation concern:

- the clear introduction of the policy dimension in the former EF;
- the unpacking of assumptions to allow a certain degree of policy analysis so as to facilitate understanding of the contextual factors that further or hinder attainment of the GBS/SBS objectives;
- the definition of a manageable number of Evaluation Questions aimed at analysing the crucial areas of the GBS/SBS effects, in a manner consistent with the EEF methodology;
- the structuring of the evaluation and of the Evaluation Questions around some major effects occurring between different levels, whereby the specific inputs that contribute to their determination are not pre-identified in a supposed causality map but are tracked through the analytical work, and their interrelations highlighted and discussed.

SECTION TWO - THE INTERVENTION LOGIC AND OTHER KEY ISSUES

Between the two approaches considered above, building an intervention logic seems the most appropriate with a view to assessing GBS/SBS as a specific aid delivery method, and checking its relevance, efficiency and effectiveness in supporting governments aiming to achieve sustainable poverty reduction. In this section, the rationale for the proposed IL will be discussed.

It should however be noted that not all forms of financial aid transfer should be assessed with the methodology proposed in the following section. There are in fact cases that fall outside the scope of the present work, where the transfer of financial aid is not linked to a framework of conditionalities aimed at supporting ex-ante the implementation of a given government programme (e.g. PRSP, or sectoral reform), but is rather seen as an ex-post reward, by one or more of a government's international partners, for the results achieved in its reform process^{9/10}. In such cases, the focus of an evaluation could be limited to an assessment of the development results attained and, for lesson learning, on the key factors (policies etc.) that have determined or show promise of determining such results.

3 THE RATIONALE FOR BUILDING AN INTERVENTION LOGIC

3.1 THE BASICS OF AN INTERVENTION LOGIC

Building an intervention logic (IL) is based on programme theory and establishes a logical sequence articulated in three basic levels (Figure 1):¹¹

- *inputs* (the means put in place by the programme)
- *outputs* (the goods and services produced by the programme through the transformation of the inputs)

⁹The emphasis on the ex-post reward is expressed in: O. Barder and N. Birdsall, *Payments for Progress: A Hands-Off Approach to Foreign Aid*, Centre for Global development - Working Paper 102 - December 2006. In this paper, the authors briefly present the experience with *output- and performance-based aid* that links i) payments to the demonstration of outputs (e.g. the fund for 'output-based aid' at the World Bank) or ii) the provision of new funding to past performance (e.g. the Global Alliance for Vaccines and Immunization and the Global Fund to Fight AIDS, TB and Malaria). From here, they then present the *progress-based aid* or payment for progress, an approach for scaling up foreign aid whereby "donors would bind themselves as a group to pay a specific amount for clear evidence of progress against one or more agreed goals [...] Payments would be determined by the achievement of the outcomes, and not linked to the implementation of any particular policies, any other intermediate outputs, or 'tied' purchases from a particular suppliers or companies."

¹⁰ In a different perspective, a similar form of financial support characterises the macro-financial assistance to third countries provided by the EC (ECFIN). In this case, the assistance of macro-economic nature, "is released on the basis of the fulfilment of economic and financial policy conditions and is provided with a view to easing the recipient countries' global external financing needs. [...] Unlike other EC financial instruments with macro-economic objectives (notably the direct budget support provided under the MEDA Regulation and the EDF Cotonou Agreement), MFA is not meant therefore to provide a regular financial support framework for structural changes or more generally the economic and social development of the beneficiary countries." For more details see EC, DG for Economic and Financial Affairs, (2008) *Vademecum on EC Macrofinancial Assistance to Third Countries*.

¹¹ Fitzpatrick, Jody. (2002). Dialogue with Stewart Donaldson. *American Journal of Evaluation*, 23(3), 347-365. Here the three levels are applied to evaluation and are defined as "actions, immediate outcomes, and long-term outcomes".

- *impacts* (the changes that such goods and services induce in the context, through interaction with the existing dynamics, and within a certain lapse of time).

Such levels are variously named according to the users, the position in the programming cycle (planning, evaluation, etc.), and the type or subject of programme. In addition, in most programmes or programme evaluations a number of intermediate levels are introduced, to better represent the priorities and foci of the relevant stakeholders, to take into account the timeframe envisaged, and to spell out the hierarchy of objectives that a given programme fits.

Therefore, if the basic criteria are respected, it is possible to build more or less complex and diversified intervention logics, using different definitions and terminologies¹². In the case of budget support evaluations, a general consideration concerns the simplicity of the schemes used. They should be as straightforward as possible and totally user-friendly, to facilitate participation by government stakeholders and decision-makers who may not be familiar with sophisticated methodological tools.

3.2 A CLOSER INTEGRATION BETWEEN GBS/SBS EVALUATION AND ASSESSMENT OF THE GOVERNMENT'S STRATEGY

When building an intervention logic (IL), the theory behind the programme should be clearly spelt out. The theory stresses the idea that GBS/SBS is not a development programme *per se*, but an aid modality that supports the development strategy of the beneficiary government. According to the theory, GBS/SBS produces a number of direct or immediate outputs which *contribute* to several changes in government financial management and policies, under certain conditions. Then such changes in government outputs should *contribute* to certain development outcomes and impacts at sectoral and comprehensive levels.

The proposed IL therefore must incorporate and spell out the anticipated contributions of GBS/SBS to the government strategy and the mechanisms through which GBS/SBS is expected to operate; at the same time it also has to represent the steps from government outputs to outcomes and impacts. Indeed, the theory does not support the linear model of causality shown in Figure 1, since it supposes two integrated logical paths: the government strategy, which provides the main framework, and the GBS/SBS, which is embedded in the former.

Before discussing how to assess the relation between GBS/SBS and the achievements of the government strategy, a better understanding of the specific IL to be built is necessary. To facilitate such understanding, a tool will be introduced in this methodological discussion that will not be used in the actual evaluations, but helps to highlight the different logical steps of the GBS/SBS evaluation process. This tool is the nested logical framework¹³, as shown in Figure 3. The nested logical framework helps respond to the question of how to combine the intervention

¹² The DAC has not standardised the definitions and terminology used in the logical framework, which is the tool generally used to represent an intervention logic. This allows a certain variety and flexibility, according to the donor agencies involved.

¹³ Interlocking or nested Logical Frameworks are developed to show how different parts of a complex program connect to each other. See Bakewell, O., and Garbutt, A. (2005), The use and abuse of the logical framework approach, SIDA or Nancholas, S. (1998) "How to do (or not to do) a Logical Framework" in Health Policy and Planning; 13(2): 189-193, & Comment and Response to the article (1999) in Health Policy and Planning; 14(1): 82-84, Oxford University Press.

logic of a GBS/SBS programme with that of the government Strategy that GBS/SBS is meant to support¹⁴.

The Government Strategy (GS) is considered here as a comprehensive programme, which incorporates a number of support programmes. GBS/SBS is considered one of such support programmes, and - in aid-dependent countries - a very important one.

Once unpacked, this support programme includes its own intervention logic and interacts at different levels with the GS in which it is embedded.

Figure 3 - A nested Framework, where GBS/SBS is embedded in the Government strategy

G/S Government's Strategy		Context and non programme factors		
Goal / Impact	Sustainable growth & Poverty Reduction	General / Sectoral Budget Support		Other sub-programmes (same BS level)
Purpose/ Outcome	Positive responses by beneficiaries - service users and economic actors - to gov. policy management and service delivery	Goal/ Impact	Contribute to the achievement of the goal/impact of the government strategy	xxx
Output	Positive changes in financing and institutional framework for public spending & public policy, and in public policy management & service delivery	Purpose/ Outcome	Contribute to strengthen the Government systems: PFM and policy processes	xxx
Input	Various Government inputs, GBS/SBS and other external assistance programmes/projects to increase funds / improve capacities	Direct outputs	1. Increased size and share of external assistance funds made available through the national budget 2. Increased size and share of budget available for discretionary spending 3. Increased predictability of external funds 4. Policy dialogue and conditionalities, coordinated, consistent with - and conducive for - government strategy 5. TA/capacity building coordinated, consistent with - and conducive for - government strategy 6. External assistance as a whole more harmonised & aligned	xxx
Context and non programme factors		Input	- Transfer of funds to budget - Policy dialogue and related conditionality - TA/capacity building - GBS/SBS aligned to government policies and systems, and harmonised across donors	xxx

¹⁴ This issue is explicitly addressed also in EC - AIDCO, DG DEV & DG RELEX, (2007) Aid Delivery Methods: Guidelines on the Programming, Design & Management of the General Budget Support, March 2007 (page 24), where it is stated that “GBS should be seen as contributing to ensuring the better functioning of an overall framework that has been established by the partner country.”

3.3 THE COMPREHENSIVE EVALUATION FRAMEWORK

Keeping in mind the conceptual framework shown in Figure 3, a simplified logical framework has been proposed, according to a more familiar model (Figure 4), for use in the evaluations. In Figure 4 the contents are the same as in Figure 3 but are represented in a different way:

- the nesting of GBS/SBS in the government strategy is represented through a simpler graphic arrangement and a heading that identifies GBS/SBS and other programmes, as “inputs to government policy and spending action”;
- the headings of the columns in Figure 4 correspond to the headings of the rows in Figure 3, with a slight difference: the GBS/SBS outputs (Figure 3) are identified as “direct outputs” (Figure 4), whereas “induced outputs” in Figure 4 are used to identify only the government strategy outputs.

This representation has been preferred because it does not emphasise the theoretical difference between the comprehensive evaluation framework and the other frameworks used in the past. The emphasis has been put on the practical evaluation method, as discussed in the following pages.

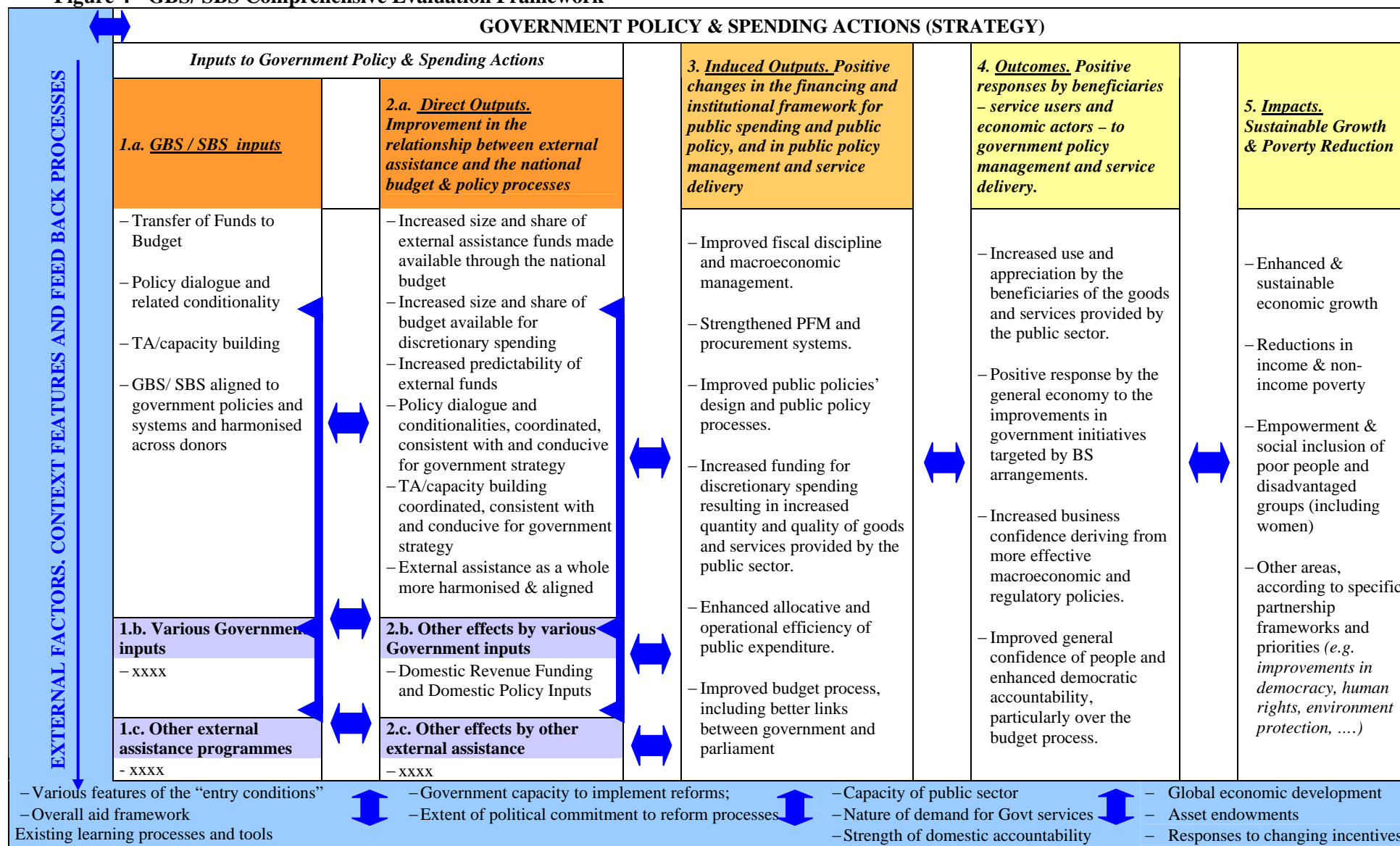
The proposed Comprehensive Evaluation Framework (CEF) represented in Figure 4 is divided into five levels included in – and interacting with – a general context shown through a row encompassing all columns. This highlights that the context is not only significant for the design of the budget support programme, but affects and is affected-by the whole sequence of effects across the evaluation framework. More specifically, the context within which GBS/SBS has to be evaluated is the overall country context in which the government strategy was conceived, has been put in place, and is being implemented; this includes the so-called *entry conditions* as well as the framework of external aid that is directly affected by GBS/SBS outputs (see Figures 3, 4 and 5).

The Evaluation Framework is divided into five levels as follows:

- Level 1:** GBS/ SBS inputs
- Level 2:** Direct Outputs (improvements expected in the relationship between external assistance and the national budget and policy processes)
- Level 3:** Induced Outputs (positive changes expected in the financing and institutional framework for public spending and public policy, and consequent improvements in public policy management and service delivery)
- Level 4:** Outcomes (positive responses by beneficiaries – service users and economic actors – to government policy management and service delivery)
- Level 5:** Impacts (sustainable economic growth, poverty reduction, including the empowerment of the poor and improvements in their real incomes, and in the case of SBS, the achievement of sector development objectives¹⁵).

¹⁵ These impacts refer mainly to sub-Saharan African countries and other impacts are to be considered according to the specific partnership framework and the related priorities established by the government and the main development partners such as significant improvements in democracy, human rights, and environment protection. With specific reference to the EC, impacts may also include: i) sustainable prosperity, stability and security (namely in the “Neighbouring Countries”); ii) the set up of a favourable environment to trade and business and sustainable stability and prosperity (namely in Latin American countries); iii) strengthened trade and investment flows and ensured development of the less prosperous countries of the region (namely in Asian countries).

Figure 4 - GBS/ SBS Comprehensive Evaluation Framework



4 THE THREE STEP APPROACH TO THE APPLICATION OF THE FRAMEWORK

4.1 AN APPROACH TO THE EVALUATION THAT FITS THE INTERVENTION LOGIC

What are the practical implications for the evaluation of the logical framework presented above?

- a first consideration is the fact that the logical framework includes two “programmes”: GBS/SBS and the government strategy, which are closely related but are not at all identical. Therefore, even if the focus of the evaluation is only budget support, an understanding of the actual implementation of government strategy is necessary;
- a second consideration regards the scope and method of such an understanding: indeed, following the linear causality process of GBS/SBS, it is logically impossible to reach a comprehensive identification of the outcomes and impacts of the government strategy and their determinant factors. The linear model leads to the identification of those outcomes and impacts of the government strategy that are strictly related to GBS/SBS inputs but does not allow the identification of the determining factors of those outcomes and impacts that are not directly related to GBS/SBS. Therefore a broad assessment of the outcomes and impacts of the government strategy is a basic pre-condition for evaluating any GBS/SBS contribution.

The previous major evaluations mentioned in our review - despite the JEGBS’ attempts to avoid the narrowness of the linear model - have tried to assess the extent to which GBS/SBS had influenced the outcomes and impacts of the government strategies. Such assessment has however been always uncertain and vague, compared to the stronger conclusions reached by these evaluations on the direct effects of GBS/SBS and the changes induced in government financial management and policy process. Indeed, the assessment of outcomes and impacts has been qualified as the weakest part of these evaluations¹⁶. The reason is that the methodology of these studies gave the highest priority to assessment of the GBS effects up to the actual changes in public expenditure. They neither allocated enough resources and time to build up the comprehensive data set, nor developed the basic policy analyses necessary to understand the level of attainment of the expected outcomes and impacts and the related determinant factors.

4.2 KEY ISSUES FOR A THREE-STEP APPROACH

Based on such considerations, the team has proposed that a ‘three-step’ approach be followed in the application of the Comprehensive Evaluation Framework (CEF). The three-step approach corresponds to the structure of the proposed IL. It explicitly recognises that budget support is a contribution to a government strategy (to a set of domestically-determined public policy and spending measures). By implication, the evaluation of GBS/SBS operations must involve *first* an assessment of the specific intervention logic of budget support (see sections 1.1, 2.1 and 3 of Figure 4), *second* an assessment of those outcomes and impacts of the government strategy that GBS/SBS intends to support (see sections 5, 4 and 3 of Figure 4), and *third* an exploration of the linkages between the GBS/SBS processes and the performance of the government strategy.

This approach aims to achieve a balance in the evaluation between the assessment of the proper effects of budget support (inputs, direct outputs and induced outputs, as defined in Figure 4) and

¹⁶ “Study teams could not confidently track distinct (separately identifiable) PGBS effects to the poverty impact level in most countries” in IDD and Associates, *Joint Evaluation of General Budget Support, Synthesis Report*, 2006, page 72. Considering the methodology adopted by the JEGBS, this conclusions applies also to what we have considered outcomes (i.e. the actual use by the target groups, of the services provided).

the assessment of the influence that GBS/SBS is able to bring to bear on the government strategy to help achieve its outcomes and impacts. At the same time it recognises the difficulty of assessing the influence of GBS/SBS at outcome and impact levels without a broader understanding of global or sectoral social and economic performance at country level. In the actual application of this approach, the scope and depth of the second step - and its relative importance *vis-à-vis* the first step - have to be decided and agreed upon before the evaluation is undertaken.

In a normal case, where the focus of the evaluation is GBS/SBS, a broad assessment of the outcomes and impacts of the government strategy may be sufficient. This implies acquisition of relevant updated data and analyses, including possible adjustments in the evaluation planning and phasing, to enable the use of up-to-date surveys and studies¹⁷. The step two exercise should at least help identify the levels of change in the key impact and outcome indicators on which GBS/SBS focus and those that may be considered the key factors of such changes. The source of information will not be primary data, but existing documentation (including quantitative data) and informed advice. The method of assessment should be the establishment of a policy matrix through a “top-down” process: from the achievements (positive or negative) to their possible determinants, highlighting any uncertainty that needs to be further investigated or any alternative view. Quick quantitative analyses may be carried out in this context, for performance assessment in specific sectors or specific areas.

In particular cases it may be agreed that a joint exercise be undertaken with the government for an overall assessment of the government strategy (namely at sectoral level, or even comprehensively) together with the evaluation of GBS/SBS. This might be requested by the government or proposed to the government by its international partners in order to maximise both the investments - in terms of expertise and time allocation - and the results, in terms of lessons learning. In such case, step two will be addressed more in depth through production - directly by the government or via different support programmes - of primary data (preparatory surveys may be planned in view of the evaluation), impact evaluations of specific government policies, and so forth.

In all cases a significant and substantial set of data must be available. This can either be produced specifically in view of the evaluation or can result from periodical annual reviews or both. In those cases where it is agreed that the production of primary data is necessary, the data gathering process should be planned with one or two years of advance to ensure the availability of reliable and processed data for the evaluation. The evaluation team is responsible for gathering and processing the data available for the evaluation, but cannot be overloaded with basic processing of primary data.

4.3 OVERVIEW OF THE THREE-STEP APPROACH

4.3.1 Key components and overlapping

The required approach combines the following:

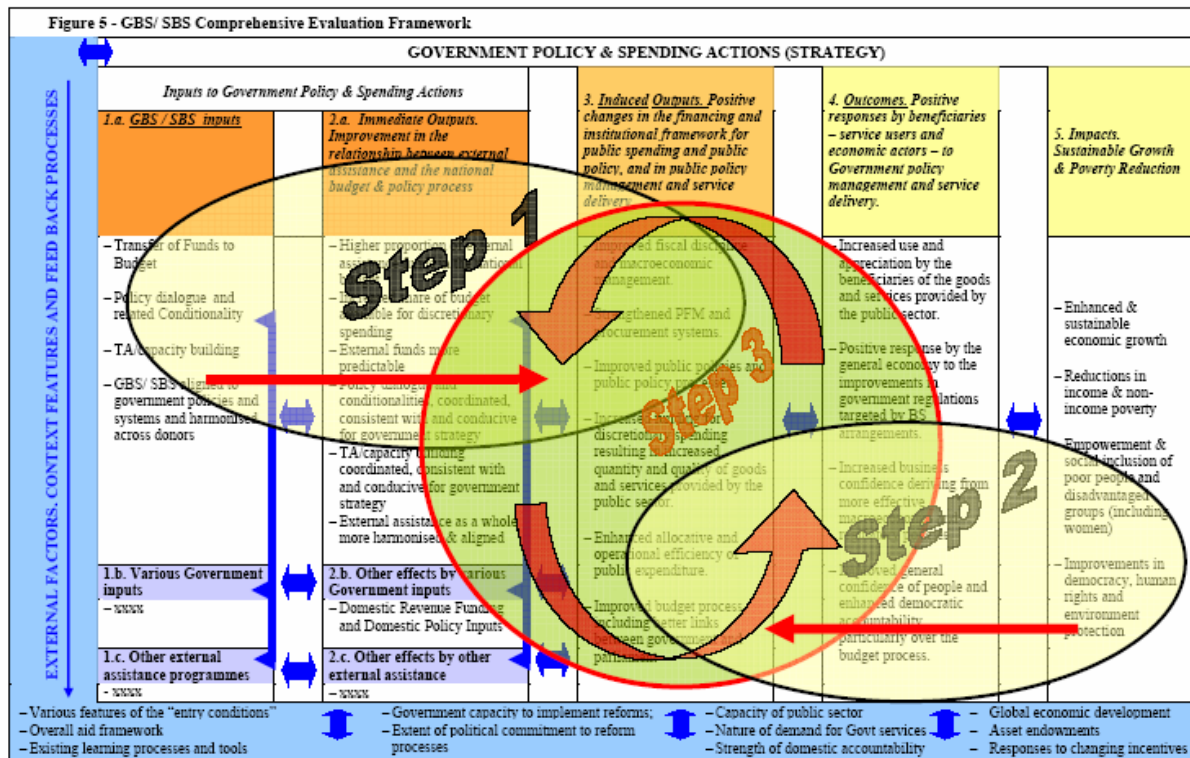
- The first step: a thorough evaluation of the budget support at levels 1, 2 and 3 of the Framework. This is focused on the relevance of the inputs provided, the direct outputs of these inputs, and the quality and adequacy of the changes supported in the government systems (induced outputs).

¹⁷ In the case of the recent evaluation in Ghana (op. cit.), a better synchronisation with the Ghana Living Standard Survey (GLSS) would have allowed the use of updated GLSS data into the evaluation.

- The second step: a broad assessment (in normal cases) of the GBS/SBS-related outcomes and impacts of the government strategies. In the normal cases the second step does not consist of an evaluation of the government strategy, but provides exhaustive information on the level of achievement of the development results that GBS/SBS supported and on the related determinant factors. In particular, it should allow the evaluators to understand:
 - whether the achievements in terms of economic growth, poverty reduction and other development objectives correspond to the GBS/SBS-supported targets of the government strategy; *and*
 - whether the main factors in possible success or failure are external to the government strategy (e.g. positive or negative external shocks) or are linked to the quality and level of its implementation, including the changes in Public Financial Management (PFM) and policy processes directly supported by GBS/SBS.
- The third step: an assessment of the contribution of GBS/SBS to the factors that have had a key role in determining the success or failure of the government strategy. It combines and compares the results of the first and second steps.

The first step is a bottom-up process covering the first three levels of the Evaluation Framework from GBS/SBS inputs to direct outputs, to the induced outputs at level 3. It should be realised that the output at level 3 is a combined result of GBS/SBS inputs and effects and government actions (see both Figures 3 and 4). The second step - the government strategy assessment - is a top-down process from the impact (level 5) and outcome (level 4) of the strategy back to its determinant factors, including the outputs of the government strategy (level 3).

Level three of the framework (see Figure 4, 'induced outputs') is part of the assessment in both Step One and Step Two. Whereas a number of government outputs, more directly related to the budget support contribution, may be easily recognised through the bottom-up approach of Step One - for example improved macroeconomic management - other outputs of government strategies will be better identified through the top-down approach of Step Two, as explanatory factors for changes at Level four, for example in public policy implementation. The limits of the bottom-up and the top-down approaches will vary depending on the specific evaluation context. The overlap will constitute the juncture between the two steps and will be important for Step Three (see Figure 5).



4.3.2 Peculiarity of Step Three

Steps One and Two are relatively traditional evaluations. From a methodological point of view, in Step One, a *process* evaluation is carried out with a focus on the changes in government systems (PFM and policy processes) induced by GBS/SBS. In step two, an *impact assessment*¹⁸ is undertaken with a focus on those outcomes and impacts of the government strategy that GBS/SBS has sought to support. The methods to be applied in these two steps are discussed in the next chapters and include building counterfactuals, combining qualitative and quantitative assessments, and so forth.

Step Three is different in nature. It compares the results of Steps One and Two and looks for internal coherence between them. As seen above, although the two steps assess - through different itineraries and different emphases - the same items included in level 3 of the Intervention Logic, their results are not necessary the same.

To take an example from the education sector, the assessments in Step One may demonstrate that there has been an improvement in the efficiency of public expenditure allocation, particularly in increasing infrastructure in poor areas and improving teachers' training. One can advance different hypotheses as to what the assessments in Step Two might demonstrate:

- there is (or is not) an actual improvement in the level of infrastructure and teachers' training: this is a consideration of the operational efficiency of public expenditure, which is very difficult to assess in Step One¹⁹. In this case Step Two provides additional information to complete the assessment of Step One.

¹⁸ Here we refer to the type of the assessment / evaluation which aims at identifying the development results of given policy variables and not to the depth of the exercise which will be decided according to the scope of the overall evaluation.

¹⁹ See also the related difficulties in JEGBS, op. cit.

- there is (or is not) a significant correlation in the target areas between such improvements in public expenditure and pupil enrolment or pupil performance (e.g. drop-out rates). The assessments in Step Two may find that the correlation between the two variables is significant in all the poor areas where infrastructure and teachers' training have been improved. Or it might find that it is too early to identify such a correlation and that the observable improvements in the outcome indicators are mostly determined by other factors, such as an increase in (say agricultural) income, an improvement in health conditions, or other.

In practice, Step Three completes and integrates the previous two steps. In our example, Step Three may help:

- complete Step One in relation to the influence of GBS/SBS on the operational efficiency of public expenditure;
- understanding of the actual contribution of GBS/SBS to pupil enrolment or performance.
- identification of new unexpected determinants of the considered outcomes that might provide additional information on GBS/SBS effects (for instance through macro-economic stabilisation, increase in agricultural commodity prices, improvements in health conditions, etc.).

Step Three is the key to completing the contribution²⁰ assessment on the causal relationship between GBS/SBS and the government strategy outcomes.

4.4 THE EVALUATION QUESTIONS

The following table shows a long list of Evaluation Questions. The list intends to cover the possible concerns and focuses of a GBS/SBS evaluation, to offer an exhaustive example for the evaluators. In the actual evaluations, the list must be adapted to (i) the policy content and intervention logic of the budget support provided to a specific country and (ii) to the political, economic and social context of that particular country. The questions must be reduced, merged or split, and reformulated according to the priorities shown by the specific Intervention Logic. Once the evaluation questions have been finalised, clear judgement criteria, indicators and targets will have to be formulated.

²⁰ It should be clear that in policy evaluations in general, and particularly in aid evaluations and in this methodology, there can be no *attribution* strictu sensu of the results of government policies to any form of aid, and the assessments are always based on a contribution analysis (search for “plausible associations”). See: Mayne, J. (1999) *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*. Discussion Paper. Ottawa: Officer of the Auditor General of Canada (page 7), and Iverson A. (2003) *Attribution and Aid Evaluation in International Development: Literature Review*, IDRC.

EF Level	Key evaluation questions
Level 1: Inputs (Step One)	1.1. Which inputs have been provided and to what extent do they correspond to the envisaged GBS/SBS inputs?
	1.2. To what extent are the budget support operations put in place relevant to the overall EC / DPs development strategies; and to the countries strategic and policy framework?
	1.3. How appropriate was the design (including the mix of GBS/SBS inputs) to the specific political, economic and institutional context?
Level 2: Direct Outputs (Step One)	2.1. To what extent has budget support contributed to an increased size and share of external funding subject to government's budgetary process and improved predictability of aid flows overall?
	2.2. To what extent has budget support contributed to a framework of policy dialogue, technical assistance and capacity building which is strategic and focused on government priorities?
	2.3. To what extent has budget support contributed to greater harmonisation and alignment of external assistance, and reduced transaction costs?
Level 3: Induced Outputs (Step One)	3.1. To what extent have there been improvements in fiscal discipline and macroeconomic management and how far has budget support contributed?
	3.2. To what extent did budget management and overall PFM improve and how far has budget support contributed to those improvements?
	3.3. To what extent have there been improvements in policy processes including the quality of policies and policy implementation, and how far has budget support contributed to those improvements?
	3.4. How has the level and composition of public spending changed and with what consequences for production of goods and services and for allocative and operational efficiency?
Levels 4 + 5: Outcomes & Impacts (Step Two)	4.1. How did the economy, in particular the enterprises, respond to the changes in the economic and institutional environment? <i>For SBS, this question will highlight the relationships between the overall economic context and sectoral performance.</i>
	4.2. How did access to services (outcome) and overall livelihoods (impact) of the target groups improve with respect to the key social sectors, in particular health and education? <i>For SBS, this question will focus on the specific sectors.</i>
	4.3. To what extent have there been improvements in governance and democratic accountability, particularly regarding the relative roles of Parliament and Civil Society in relation to the Budget? ²¹
	4.4. To what extent did the rule of law improve in the country, with particular respect to citizens' security and the access of entrepreneurs and the poor to justice?

²¹ It is postulated that both political and administrative accountability can be reinforced through the dissemination of information, and by the advocacy and mobilisation of civil society groups. Within this framework, some donors, and especially donors engaged in GBS operations, are pursuing anticorruption strategies and actions aimed at increasing political and administrative accountability by other, complementary means, including the provision of TA to support accountability institutions (audit agencies, parliaments etc), and support to civil society organisations engaged in such issues.

	4.5. To what extent can significant improvements be identified in some key cross-cutting themes in the society, such as environment protection, gender equality and youth participation?
Level 3: Induced outputs and other determinants of key outcomes (Step Two)	5.1. To what extent was economic success or failure related to changes in macroeconomic and fiscal management and other economic policies put in place by government, or to other external or internal factors?
	5.2. To what extent was success or failure in increased access to services by and improved livelihoods for the poor related to changes in allocative and operational efficiency of public expenditure, in government policy processes, or in other external or internal factors?
	5.3. To what extent was success or failure in improving economic governance and transparency and access of the poor and entrepreneurs to justice related to changes in government policy processes, and/or in other external/ internal factors?
	5.4. To what extent was success or failure in improving any specific outcomes in some key cross-cutting themes related to changes in allocative and operational efficiency in public expenditure, in government policy processes, or in other external or internal factors?

Step Three: rationale for the formulation of the evaluation questions	<p>The questions should be formulated on completion of the first two steps, based on the coherence and complementarity of, or discrepancies in, the answers provided in Steps One and Two. They should aim at highlighting any significant causal link (contribution) between GBS/SBS and the actual outcomes identified. The rationale for the formulation of the EQs may be defined as follows:</p> <ul style="list-style-type: none"> ▪ Step One general question asks: to what extent did GBS/SBS contribute to the achievement of the government outputs in level 3 of the IL? ▪ Step Two general question asks: to what extent the achievement of the development outcomes that GBS/SBS aimed to support has been favoured, hampered, or basically not influenced by the government outputs in level 3 of the IL. ▪ Step Three general question asks: to what extent were the GBS/SBS contributions to government outputs consistent with those outputs' contribution to the intended development outcomes?
---	--

4.5 ADJUSTING THE FOCUS OF THE EVALUATION TO THE COUNTRY'S LEVEL OF AID DEPENDENCY

In general, there are two significantly different contexts in which budget support tends to be provided:

- an aid-dependent country where the combination of GBS and SBS as a percentage of government expenditure reaches double-digit levels, as in most Sub-Saharan African countries receiving budget support;
- a non-aid-dependent country (often lower-middle-income country) in which budget support is typically provided as part of a wider political partnership in order to support progress along a well-defined reform path to which the government has subscribed

through political agreements / treaties, for example the association process in the EU's pre-accession and neighbouring countries²².

In an aid-dependent country, budget support has a quantitative impact that clearly makes a difference. In this context the GBS/SBS inputs provided are deliberately intended to affect specific steps in the policy and public spending process and to trigger virtuous sequences of effects. Clearly in this context there is a specific interest in assessing whether and how the inputs fulfilled the intended role and in tracking the effects of the funds and other inputs provided. In such a context, it is likely that evaluation managers will want to put substantial emphasis on Step One of the evaluation approach and will want each of the recommended Evaluation Questions addressed.

In the case of non-aid-dependent countries, budget support is likely to represent a much lower percentage of government expenditure (1% or less in many transition countries). Hence the identification of specific flow of funds effects and of changes in institutional and policy processes linked to specific inputs provided through a GBS or SBS aid package may be more difficult. In such a context, the budget support inputs are generally intended to help the governments overcome specific budget constraints and provide generic and general support (*ex ante* encouragement, or *ex post* rewards) to strengthen given policies or programmes. Often in such contexts, budget support is a part or a complement of huge capacity building programmes, including approximation of the legislation, political and institutional reform, support to the civil society. It would be very difficult to isolate the specific effects of GBS/SBS. Here, it is possible that neither the recipient government nor its development partners will be interested in focusing on the specific mechanisms and causal links through which budget support may have worked. In these cases, it is likely that the interest will mainly focus on assessing the achievement of the intended results and thus on the success or failure of the Government strategy and possibly its main drivers, for the purposes of lesson-learning.

Here the relative emphasis of the evaluation is likely to be on Step Two and there may be an interest in developing more detailed Evaluation Questions to address specific aspects of the government strategy being assessed. By contrast, Evaluation Questions concerning the direct outputs budget support might be of less interest.

Even where the scale of budget support is small, it is absolutely legitimate to suppose that budget support creates a partnership framework that facilitates government access to much needed and diverse know-how in support of its policy process, and that the marginal availability of a small amount of funds may allow the government to undertake specific development initiatives that might have been postponed in the absence of such funds. Thus Evaluation Questions relating to the nature and timing of GBS/SBS inputs will remain important. Similarly, Step Three of the evaluation, focusing on the relationship of budget support to the achievement of government's outcomes and impacts, will also remain relevant. However, it is important to retain a sense of pragmatism. It will always be difficult to evaluate the marginal effects of GBS/SBS on a process driven and determined by mostly endogenous, politically-driven and decidedly stronger factors. The key question will always be whether the government strategy is working or not. Moreover, there is the implicit consideration that the partner government is politically committed and, as such, evaluations should focus on helping the government improve the effectiveness of its strategy.

²² An example is the budget support provided to North African countries under the MEDA instrument and now the ENPI instrument.

4.6 ADJUSTING THE FOCUS OF THE EVALUATION TO THE TYPE OF BUDGET SUPPORT

The proposed intervention logic, approach and EQs may be applied to three different evaluation contexts:

- situations where only General Budget Support is provided;
- situations where General Budget Support is provided alongside one or more Sector Budget Support arrangements;
- situations where Sector Budget Support is provided on its own.

The recommended Evaluation Questions presented in §4.4 are likely to be relevant for all three scenarios, although in the case of SBS there will be a need to deliberately apply the relevant questions both at general level and to the sector or sectors where SBS is being provided. For example in relation to direct outputs, one would need to ask whether there was evidence of increased harmonisation and alignment of development assistance, both at aggregate level and within the sector or sectors receiving SBS. This dual application of the Evaluation Questions would be relevant for all questions except those referring to macro-economic outputs, which by definition cannot be examined from a sectoral perspective.

Even where Sector Budget Support is provided without an accompanying GBS arrangement, it would normally be appropriate to address the Evaluation Questions at both sectoral and national levels. There are two reasons for this. In the first place there are many SBS arrangements which are of a sufficient scale to generate a noticeable impact at national level, for example on the overall availability of discretionary resources within the national budget. In some cases, the creation of a 'demonstration effect' outside the targeted sector may be a deliberate objective. For example procurement reforms may be deliberately pursued within a Health Budget Support arrangement as a way of contributing to the definition of a national procurement strategy.

Second, it will be important to track the most significant changes taking place at national level in order to understand what is driving the changes at sectoral level and in order to assess the causal effects of SBS to a reasonable degree of accuracy. For example, a modest increase in sector funding may be a significant achievement in the context of a constrained macro-economic situation and might suggest an important contribution from the SBS arrangement. The assessment would be quite different, however, in a context of fast growth in domestic revenues. Similarly, improvements introduced in the quality of parliamentary scrutiny of public spending may be of some significance at sectoral level but could not realistically be attributed to SBS.

Besides the need to apply most standard Evaluation Questions at both aggregate and sector level, some aspects of SBS demand specific consideration and emphasis. Key issues specifically related to SBS include for example the sector policy focus, sector level outputs targeted in the SBS arrangement, performance measurement at sector level, adequate funding of the sector budgets, government-donor relations at sector level, alignment (e.g. SBS versus basket funds, procurement issues), and so on. Such issues may be highlighted through the adjustment of the standard questions, or the formulation of new questions (in addition or replacement of the existing ones).

Regarding Step Two, SBS evaluation provides an opportunity for deeper and more sophisticated assessments in terms of the outcomes achieved and their determinant factors. Quantitative analyses (see chapter below) and impact evaluation of specific sectoral policies might be developed together with the standard qualitative policy matrix.

5 COUNTERFACTUALS AND OTHER EVALUATION TOOLS

5.1 COUNTERFACTUALS

5.1.1 *Counterfactual analysis and impact evaluation*

Counterfactual considerations are implicit in the application of any evaluation criterion. Indeed, all evaluation judgements - whether concerning relevance, efficiency, effectiveness, impact, and so on - entail the use of an implicit measurement unit (or a term of comparison) that must be clearly explained²³.

The DAC Glossary of Evaluation Terms (*OECD DAC 2002*) defines the counterfactual as: “*the situation or condition, which hypothetically may prevail for individuals, organisations or groups were there no development intervention.*” Therefore “the counterfactual, or counterfactual scenario, is an estimate of what would have occurred in the absence of the evaluated intervention”, and the purpose of the counterfactual is to assess the effect of the intervention by subtracting the counterfactual from the observed change (factual)²⁴. Notwithstanding that the no-intervention situation is the standard alternative, whenever possible the examination of the best (or most realistic) alternative may be much more interesting for the policy-maker²⁵.

The practical issue, with a number of methodological implications is how to build credible counterfactuals in the specific case of GBS/SBS evaluations.

5.1.2 *Different approaches to counterfactuals*

Three different approaches to building a counterfactual in an evaluation are considered:

- i. Quantitative models: these consist of building a quantitative model of the macro or sectoral context in which the programme under evaluation operates, and through which it is possible to test the impact of different policies (or no policies at all), compared to those under evaluation. It provides the opportunity for a simulated counterfactual.
- ii. Control groups: these are based on considering as counterfactual what happened, during the same period of the programme under evaluation, in other areas of the same country (or other countries) that present strong similarities to the area or country under evaluation, but where different policies (or no policies at all) have been implemented;

²³ Many authors consider counterfactual analysis as specific to impact evaluation. M. Ravallion in his “Evaluating Anti-poverty Programs” (World Bank Policy Research Working Paper 3625, June 2005) states that “Impact evaluation (or “counterfactual analysis”) assesses outcomes for a specific program”; while the DAC Network on Development Evaluation refers to four most common meanings of impact evaluation, among them: “i) An evaluation which looks at the impact of an intervention on final development outcomes, rather than only at project outputs, or a *process evaluation* which focuses on implementation; ii) An evaluation concerned with establishing the counterfactual, i.e. the difference the project made (how indicators behaved with the project compared to how they would have been without it);”. See *Impact Evaluation: An Overview and Some Issues for Discussion*. Independent Evaluation Group of the World Bank, in collaboration with the DAC Secretariat. The World Bank Poverty Net website emphasis the importance of counterfactual analysis (second of the four DAC definitions), “The central impact evaluation question is what would have happened to those receiving the intervention if they had not in fact received the program. Since we cannot observe this group both with and without the intervention, the key challenge is to develop a counterfactual”.

²⁴ European Commission, EuropeAid (2006), Evaluation Methods for the European Union’s External Assistance: Methodological Bases for Evaluation. EuropeAid website: http://ec.europa.eu/europeaid/evaluation/methodology/methods/mth_cft_en.htm

²⁵ Ravallion M. (2005) *Evaluating Anti-poverty Programs*, World Bank Policy Research Working Paper 3625.

- iii. **Qualitative alternative scenarios:** these consist of mainly qualitative alternative scenarios built as a counterfactual to the actual context under evaluation. Such scenarios are built using theory, and lessons from past experience in the same country or other similar countries. They include alternative policies or instruments, or no policies at all, compared to the situation under evaluation.

The selection of one of such methods depends upon a number of factors, two of which are the most important: the nature of the programme under evaluation, which may allow or not allow the use of one or more of such approaches; and the resources available, since some of these approaches need specific skills, data and time which may not be compatible with the scope of a GBS/SBS country evaluation.

Quantitative models

An exhaustive presentation and discussion of the use of quantitative models to test the impact of macro and sectoral economic policies on the various groups of the relevant population is provided in Bourguignon and Pereira da Silva *The impact of Economic Policies on Poverty and Income Distribution (Evaluation techniques and tools)*, OUP 2003.

The focus of such techniques is normally on the evaluation of the policy impact and not on the evaluation of aid effectiveness. Since most GBS/SBS evaluations so far have related to the evaluation of aid, such techniques have not been adequately considered. In the proposed methodology (three step approach), where much room is provided for policy impact assessment, the use of such techniques may be reconsidered, including within the evaluation design adequate skills, financial resources and time. The models proposed are micro and macro models, with the inclusion of a new generation of micro-macro linkage models.

The micro models are mainly suitable for specific or sectoral policy evaluations. They may be used to test how shifts in public expenditure in a given field or sector may affect the welfare of the target population and other groups involved. Indeed, the core of such techniques is the poverty incidence analysis, which allows assessment of whether and in what measure the groups targeted by the intervention (or other interrelated groups) have benefited from it. The data necessary for such applications are relatively limited, including up-to-date household surveys, possibly disaggregated at geographical level. The software in most cases is a standard statistical package. The macroeconomic tools require aggregated or disaggregated economy-wide data. Sometimes the basic data used by IMF and WB for the country financial framework are sufficient²⁶. At other times the data required are more sophisticated²⁷.

All such tools may be used both for *ex post* and *ex ante* evaluation. They are particularly suitable in the preparation of PRSP or government sectoral programmes, as well as in those *ex post* evaluations where particular importance is given to the assessment of the government policies supported by GBS/SBS. This may be the case in evaluations in which the government has strong participation: as verified by actual experience, governments are more interested in capturing the impact of the policies implemented, rather than in analysing the way in which aid has affected such policies.

²⁶ This is the case in Shantayanan Devarajan and Delfin S. GO, *The 123PRSP Model*, Chapter 13 of F. Bourguignon, L. A. Pereira da Silva, *op. cit.*

²⁷ Ex. Jeffrey Round, *Social Accounting Matrices and SAM-based Multiplier Analysis*, Chapter 14 of F. Bourguignon, L. A. Pereira da Silva, *op. cit.*

Control groups

This approach is one of the most used and sometimes is considered the only really valuable approach to counterfactuals²⁸. The control group approach can provide valid information when systematic differences between the two individuals, groups or areas can be ruled out or accounted for. In particular, two sets of problems must be addressed:

- i. The selection bias: that is, when the beneficiaries of the intervention are selected on the basis of (to the evaluator) unobserved characteristics that influence outcomes, the comparison group will not necessarily present the same set of characteristics and therefore the observed differences in measured outcomes between the units (participating and control group) cannot be entirely attributed to the programme.
- ii. The contagion bias: that is, the comparison group may be contaminated either by spill-over effects from the intervention itself or a similar intervention being undertaken in the comparison area by another agency.

In order to avoid such biases, experimental design - that is when a programme is designed with a focus on identification of beneficiary groups and control groups according to strictly observable similar characteristics - is the ideal situation. In the absence of an experimental design, a careful *ex post* identification of the control group is necessary. Indeed, this method has been tested for evaluation of specific programmes, targeted on certain observational units (be they individuals, social groups, villages, areas) and excluding other comparable units in similar circumstances. None of these conditions, however, applies in most GBS/SBS programmes and in all cases where evaluation of complex macro or sectoral policies is involved. "In most [such] cases, the intervention is *not discrete and not necessarily targeted at a specific group*". Therefore, it will be difficult or impossible to discern a treatment group and a control group"²⁹.

The move from project aid to programme aid and budget support raises the issue of assessing the effectiveness of sector or general budget support, and the related policy impacts. These are forms of aid - supporting government strategies - that are concerned with the impact of numerous heterogeneous interventions and the related inter-relationships. One of the major objectives in evaluating GBS/SBS is to catch the systemic impact of the interventions either at sector or comprehensive level (support to sector-wide or global poverty reduction and growth strategies). In these cases a control group approach may be applied as a complementary method at micro level, for instance to compare villages that benefited from some of the measures provided in a given sector (education, health, local development funds...) with villages which are in the intervention pipeline but have not yet been reached.

²⁸ Esther Duflo and Michael Kremer, *Use of Randomization in the Evaluation of Development Effectiveness*, Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness in Washington, D.C. 15-16 July, 2003. See also: White, H. (2006) *Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank*, the World Bank, Washington D.C., http://mpr.ub.uni-muenchen.de/1111/1/MPRA_paper_1111.pdf; *When Will We ever learn? Improving Lives through Impact Evaluation*, report of the Evaluation Working Group, Centre for Global Development, May 2006; *Impact Evaluation: An Overview and Some Issues for Discussion*, note prepared by the Independent Evaluation Group of the World Bank, in collaboration with the DAC Secretariat, for discussion at the 4th meeting of the DAC Network on Development Evaluation, 30 – 31 March 2006. M. Ravallion "Evaluating Anti-poverty Programs", World Bank Policy Research Working Paper 3625, June 2005.

²⁹ NONIE Sub-group 3 *Impact evaluation of new aid instruments and country programs* – May 17 - 2007. The italics is in the text.

Alternative scenarios

A pragmatic and flexible approach to building counterfactuals consists of identifying alternative ways (e.g. different policies and measures) of achieving the same outcomes pursued by GBS/SBS and the related government strategies.

The most common alternative scenario - particularly suitable for Step One of the proposed methodology - is represented by project aid *vis-à-vis* GBS/SBS. Based on past experience or experience in similar countries, and on qualified informed advice, the evaluator tries to determine what could have happened under a different aid scenario. Such an approach may be applied also in Step Two of the proposed methodology, that is to ascertain whether a given outcome would have been more or less marked under an alternative sector or macro policy.

The disadvantage of this approach is its low level of rigour, especially when the assessments are mainly made on qualitative data. The big advantage however is its flexibility and the possibility of building specific counterfactuals for almost every Evaluation Question, to guide the analyses of the evaluator and submit its conclusion to specific checks.

This approach has been used in most GBS evaluations so far, including the JEGBS (2006)³⁰ and has proved useful even though sometimes its application risks appearing superficial. Indeed this approach should be used to remind the evaluator of what could have been done (or not done) instead of what has actually been done. It may help to formulate the right questions and develop the correct approach to addressing them.

It should not be intended, however, as a way of avoiding more rigorous analyses, including - whenever possible - observation of control groups, statistical analyses, possible modelling, and so on. The resource constraints, beyond the suitability of the different methods discussed so far, should be considered a key determinant of which combinations of approaches and methods are feasible. On the question of such combinations it is worth noting that the European Evaluation Society (EES) has recently issued a statement^{31/32} in which, while welcoming the current interest in improving impact evaluation and assessment with respect to development and development aid, it deplores one perspective strongly advocated at present: that the best or only rigorous and scientific way of doing impact evaluations is through randomised controlled trials (RCTs). In contrast, the EES supports multi-method approaches to impact evaluation while promoting new research on the subject.

5.2 QUANTITATIVE AND QUALITATIVE METHODS

5.2.1 Mixed methods

A strong recommendation along the lines of mixed methods in impact evaluation entails the combined use of quantitative and qualitative methods. “A mixed-method approach requires an integrated strategy in which the strengths and limitations of quantitative and qualitative methods are recognised and the evaluation is designed to take advantage of the complementarities between

³⁰ JEGBS cit. See annex 3 for a synthetic presentation

³¹ EES Statement: The importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions. Nijkerk, December 2007. Concerns on the use of randomised trials have also been expressed by the EC.

³² Concerns on the use of randomised trials have also been expressed by the EC.

the different methods”³³.

Such a combination has not been a main concern in GBS/SBS evaluations to date. White shows how participatory methods and qualitative observations may be successfully combined with quantitative comparisons of groups, or with statistical analyses of panel data³⁴.

5.2.2 *Statistical analyses*

Various statistical analyses may be carried out in GBS/SBS evaluations to measure the factors that are more strongly associated with the achievement or non-achievement of the intended outcomes. These methods are particularly appropriate in Step Two of the proposed methodology, where the assessment aims at identifying the specific impact of different policy and non-policy factors.

In a recent study³⁵, the changes of an outcome variable in the education sector - for example the exam scores of pupils - are statistically correlated with changes in several policy variables (availability of textbooks, number of classrooms, level of training of teachers or school heads, facilities, etc.). In SBS evaluations this method may provide some interesting results *vis-à-vis* understanding of the policy processes that brought about (or failed to bring about) the attainment of the desired outcome(s), specifically, when combined with the alternative scenarios approach and qualitative observations. Indeed an exclusive use of this method would over-simplify the analysis, overlooking the cross-effects and respective weights of the different policy variables as well as the effects of the systemic changes. On the other hand, however, although the method attempts to give an account of the context-related factors, the impact of non-policy variables needs to be better highlighted.

Other statistical methods may also be applied. In particular, multivariate techniques (principal components, cluster, correspondence analysis) seem particularly suitable for identifying the most influential factors in the achievement of given outcomes at sectoral or global level. Indeed, the state of the art in GBS/SBS evaluation suggests strong encouragement toward an increased use of statistical tools. But considering the limits of the application of such tools (mainly due to poor data), they should always be combined in an appropriate mix of qualitative and quantitative assessments.

³³ Bamberger, M., Rugh, J., and Mabry, L., “Mixed-Method Evaluation”, Chapter 13 of *Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints*, Sage Publications, 2006

³⁴ Howard White, *Of Probabilities and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation* - NONIE Working Paper No. 6 - January 2008

³⁵ Elders, C., Gunning, J.W., and de Hoop, K., (2007) *Assessing Budget Support with Statistical Impact Evaluation*, VU University of Amsterdam, Tinbergen Institute, and AIID. <http://www.tinbergen.nl/discussionpapers/07075.pdf> See annex 3 for a synthetic presentation.

ANNEX 1: BIBLIOGRAPHY

- Acevedo L., Christensen, M., (2004) *Development of Swedish General Budget Support 1990–2003*, SIDA
- Aiyar, S., Ruthbah, U., (2008), *Where Did all the Aid Go? An Empirical Analysis of Absorption and Spending*, IMF Working Paper WP/08/34
- Baker, J. (2000) *Evaluating Impact of Development Projects on Poverty, A Handbook for Practitioners*, the World Bank, Washington D.C.
- Bakewell, O., and Garbutt, A. (2005), *The use and abuse of the logical framework approach*, SIDA
- Bamberger, M., Rugh, J., Mabry, L.. (2006) *RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints*, by Sage Publications Inc.
- Barder, O., Birdsall, N. (2006) *Payments for Progress: A Hands-Off Approach to Foreign Aid*, Centre for Global development - Working Paper 102
- Bourguignon, F. and Pereira da Silva, L. A., eds. (2003) *The impact of Economic Policies on Poverty and Income Distribution (Evaluation techniques and tools)*, OUP.
- Centre for Global Development – CGD (2006) *When Will We ever learn? Improving Lives through Impact Evaluation*, report of the Evaluation Gap Working Group
- Chen, H. (1990). *Theory-Driven Evaluations*. Newbury Park, CA: Sage Publications
- CIDA, (2005) *Operational Guide on Direct Budget Support and Pooled Funding to Recipient Countries*
- Cordella, T., Dell'Ariccia, G. *Budget Support vs. Project Aid*, IMF Working Paper 03/88. <http://www.imf.org/external/pubs/ft/wp/2003/wp0388.pdf>
- Dollar, D. et al. (1998). *Assessing Aid: What Works, What Doesn't and Why*. New York: Oxford University Press for World Bank
- DANIDA, (2006) *Principles for provision of General Budget support*
- DFID, (2004) *Poverty Reduction Budget Support, Policy paper*
- Driscoll, R., Christiansen, K., Booth, D with de Renzio, P., Smith, S. and Herneryd, K., (2005). *Progress reviews and performance assessment in poverty-reduction strategies and budget support A survey of current thinking and practice*, Report submitted to the Japan International Cooperation Agency, Overseas Development Institute
- Duflo, E. and Kremer, M., *Use of Randomization in the Evaluation of Development Effectiveness*, Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness in Washington, D.C. 15-16 July, 2003.
- Elders, C., Gunning, J.W., and de Hoop, K., (2007) *Assessing Budget Support with Statistical Impact Evaluation*, VU University of Amsterdam, Tinbergen Institute, and AIID. <http://www.tinbergen.nl/discussionpapers/07075.pdf>
- European Commission, AIDCO, DG DEV & DG RELEX, (2007) *Aid Delivery Methods: Guidelines on the Programming, Design & Management of the General Budget Support*.
- European Commission, AIDCO (2004) *Aid Delivery Methods - Project Cycle Management Guidelines*,
- European Commission, DG Development (2005). *EC Budget Support: An Innovative Approach to Conditionality*
- European Commission (2005). *Institutional Assessment and Capacity Development: Why, what and how?* September 2005, Brussels: European Commission.
- European Commission (2005). *The European Consensus on Development*. Joint Statement by the Council

and the Representatives of the Governments of the Member States meeting within the Council, the European Parliament and the Commission

European Commission, (2006), Evaluation Methods for the European Union's External Assistance: Methodological Bases for Evaluation. EuropeAid website: http://ec.europa.eu/europeaid/evaluation/methodology/methods/mth_cft_en.htm

European Commission, DG DEV (2006) "European Commission Budget Support: An Innovative Approach to Conditionality", in *Budget support as more effective aid? Recent Experiences and Emerging Lessons*, Koeberle S., Stavreski Z., and Walliser J. (eds), The World Bank, Wash. DC.

European Commission, EuropeAid (2007) Tools and methods Series, Guidelines No. 2 - Support to Sector Programmes: Covering the three financing modalities: Sector Budget Support, Pool Funding, and EC project procedures

European Commission, AIDCO, DG DEV & DG RELEX (2007) Aid Delivery Methods: Guidelines on the Programming, Design & Management of General Budget Support

European Court of Auditors (2005). Information note by the European Court of Auditors on Special Report No 2/2005 concerning EDF budget aid to ACP countries: the Commission's management of the public finance reform aspect. European Court of Auditors: Luxembourg.

European Evaluation Society (December 2007), EES Statement: The importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions. <http://www.europeanevaluation.org/news?newsId=1969406>

Factora, J. (2006). "A Stocktaking of Poverty Reduction Support Credits", in *Budget Support as More Effective Aid? Recent Experiences and Emerging Lessons*, S. Koeberle, Z. Stavreski, J. Walliser (eds), The World Bank, Washington D.C.

Fitzpatrick, J. (2002). "Dialogue with Stewart Donaldson". *American Journal of Evaluation*, 23(3), 347-365

Foster, M., Killick, T. (2006), What Would Doubling of Aid do for Macroeconomic Management in Africa?, ODI Working Paper 264

IDD and Associates (2005, 2006, 2007) *Joint Evaluation of General Budget Support 1994–2004*. Inception Report, Synthesis Report, Note on Approach and Methods

IDD and Associates (2006). *Joint Evaluation of General Budget Support 1994–2004*:

Independent Evaluation Group of the World Bank (2006) *Impact Evaluation: An Overview and Some Issues for Discussion*, Room Document n. 5, DAC Network on Development Evaluation <http://www.cgdev.org/doc/eval%20gap/HowardWhiteDAC.pdf>

IMF (2005), *The Macroeconomics of Managing Aid Inflows: Experiences of Low-Income Countries and Policy Implications*. <http://www.imf.org/external/np/pp/eng/2005/080805a.htm>.

Iverson A. (2003) *Attribution and Aid Evaluation in International Development: Literature Review*, IDRC.

Koeberle S., Stavreski Z. (2005), *Budget Support: Concept and Issues*, The World Bank, Wash. DC.

Lawson, A. and Booth, D. (2004). *Evaluation Framework for General Budget Support* commissioned on behalf of the OECD DAC Evaluation Network. London: ODI

Mayne, J. (1999) Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly. Discussion Paper. Ottawa: Officer of the Auditor General of Canada

Mosley P., Harrigan J., Toye J. (1995), *Aid and Power, The World Bank and Policy-based Lending*, Routledge - London and New York

- Nancholas, S. (1998) “How to do (or not to do) a Logical Framework” in *Health Policy and Planning*; 13(2): 189-193, & Comment and Response to the article (1999) in *Health Policy and Planning*; 14(1): 82-84, Oxford University Press
- Nilsson, M. (2004). *Effects of Budget Support – A Discussion of Early Evidence*. UTV Working Paper 2004:4. Stockholm: SIDA
- NORAD, (2006) Donor Definitions of and practices in providing budget support with particular reference to sector budget support Discussion Report
- ODI and Ghana Centre for Democratic Development (2007). Joint Evaluation of Multi-Donor Budget Support to Ghana
- OECD (2005), *Paris Declaration on Aid Effectiveness: Ownership, Harmonisation, Alignment, Results and Mutual Accountability* - Paris, <http://www.oecd.org/dataoecd/11/41/34428351.pdf>
- OECD/DAC (2006). Harmonizing Donor Practice for Effective Aid Delivery: Volume 2; Budget Support, Sector-Wide Approaches and Capacity Development in Public Financial Management.
- OECD DAC (2002) *DAC Glossary of Evaluation Terms*
- PEFA Secretariat (2005) Public Expenditure and Financial Accountability – Performance Measurement Framework), The World Bank, Washington D.C.
- Ravaillon, M. (2005) *Evaluating Anti-poverty Programs*, World Bank Policy Research Working Paper 3625
- Strategic Partnership for Africa. (2005), Sector Budget Support: A Note from the Dublin Workshop of SPA Working Groups, 5-6 October 2005
- Strategic Partnership for Africa – Budget Support Working Group (2005). Survey of the Alignment of Budget Support and Balance of Payments Support with National PRS Processes
- Tabor, S. R. (2005) *Good Practices for the Evaluation of Policy-Based Lending by Multilateral Development Banks?*. Report prepared for the Evaluation Cooperation Group of the Multilateral Development Banks
- Third International Roundtable on Managing for Development Results (2007), *Mutual Accountability: The Key Driver for Better Results, A Background Paper* – Vietnam, February 5-8, 2007
- USAID (2005) General Budget Support: Key findings of Five USAID studies, PPC Evaluation Paper No.7
- White, H. (2006) *Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank*, World Bank, Washington D.C., http://mpr.ub.uni-muenchen.de/1111/1/MPRA_paper_1111.pdf
- White, H., (2008) Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation - NONIE Working Paper No. 6
- White, H. and Toye, J. (eds). (1996) “Evaluating Programme Aid”, *IDS Bulletin Vol. 27 n. 4*

ANNEX 2: SUMMARY OF THE MAIN RECENT DEFINITIONS OF BUDGET SUPPORT

Institution and Source	Definition
<p>European Commission: Aid Delivery Methods: Guidelines on the Programming, Design & Management of General Budget Support, AIDCO DEV RELEX, January 2007</p>	<p>Budget support is the transfer of financial resources of an external financing agency to the National Treasury of a partner country, following the respect by the latter of agreed conditions for payment. The financial resources thus received are part of the global resources of the partner country, and consequently used in accordance with the public financial management system of the partner country.</p> <p>More holistic view of budget support “provision of money on the basis of the fulfilment of conditions in a context of dialogue, harmonisation, alignment, and capacity development.”</p>
<p>OECD/DAC Harmonizing Donor Practice for Effective Aid Delivery: Volume 2; Budget Support, Sector-Wide Approaches and Capacity Development in Public Financial Management”, 2006</p>	<p>Budget support is defined as a method of financing a partner country’s budget through a transfer of resources from an external financing agency to the partner government’s national treasury. The funds thus transferred are managed in accordance with the recipient’s budgetary procedures. [.....] In the case of <u>general budget support</u>, the dialogue between donors and partner governments focuses on overall policy and budget priorities, whereas for <u>sector budget support</u> the focus is on sector-specific concerns.</p>
<p>CIDA, Operational Guide on Direct Budget Support and Pooled Funding to Recipient Countries, 2005.</p>	<p>Direct budget support is provided directly to host-country institutions to be spent as part of their overall program budget using their own financial management systems.</p> <p>General Budget Support covers financial assistance as a contribution to a host government's overall national budget in support of a general strategy and medium-term budget agreed between the donor and host-country government. Sector budget support covers financial aid earmarked for a discrete sector or sectors, with any conditionality relating to these sectors.</p>
<p>DANIDA, Principles for provision of General Budget support May 2006</p>	<p>General budget support is an assistance package comprising several forms of input. In addition to financial transfers to the recipient country’s Treasury for co-financing the poverty reduction strategy, the package includes dialogue on policy; agreement on conditions regarding reforms and allocation of resources; mechanisms for improving the effectiveness of the total package of development assistance; monitoring – of progress towards reaching agreed goals and of the extent to which public expenditures benefit the poorest – as well as support for capacity building aimed at creating a more efficient public administration; and anti-corruption.</p>
<p>DFID, Poverty Reduction Budget Support, Policy paper, 2004</p>	<p>Poverty Reduction Budget Support (PRBS) as a form of financial aid in which funds are provided: (i) in support of a government programme typically focusing on growth, poverty reduction, fiscal adjustment and strengthening institutions, especially budgetary processes; and (ii) directly to a partner government’s central exchequer, to spend using its own financial management, procurement and accountability systems.</p>

<p>NORAD, Donor Definitions of and practices in providing budget support with particular reference to sector budget support Discussion Report, 2006</p>	<p>Budget support is a term used for aid funds that are managed by the partner country's financial system for funding of the state budget, either as a general contribution to the partner government state budget or for financing specific sectors</p>
<p>SIDA, Development of Swedish General Budget Support 1990–2003, Lorena Acevedo Narea Martin Christensen, 2004</p>	<p>Budget support is “financial means provided to the government budget's revenue side”. This support is linked to a reform programme and/or implementation of the recipient country's poverty reduction strategy. There are two forms of budget support: GBS and sector programme support. GBS refers to non-sector specific (applies to both distribution of financial resources and policy dialogue), while sector programme support is earmarked for a specific sector. GBS does not necessarily have to be given to the national budget, it can also be provided at the regional level.</p>
<p>WB; A Stocktaking of Poverty Reduction Support Credits (PRSCs): Factora, J., in <i>Budget Support as More Effective Aid? Recent Experiences and Emerging Lessons</i>, S. Koeberle, Z. Stavreski, J. Walliser (eds), 2006</p>	<p>Budget support is commonly understood to mean the provision of financial assistance directly to a partner country's budget on a regular basis, using the country's own financial management systems and budget procedures. <i>Poverty Reduction Support Credits [embody the Bank's contribution to the emerging use of budget support] typically consist of three or four annual single-tranche operations, phased to support the government's medium-term development objectives</i></p>

ANNEX 3: USE OF COUNTERFACTUALS IN THE EVALUATION OF COMPLEX PROGRAMMES

Approach to Counterfactuals in the JEGBS country assessments*

Following the statement whereby “Any evaluation requires consideration of the counterfactual, and the evaluation will be meaningful only if the counterfactual is a relevant one” the report sets out the approach to counterfactuals adopted during the JEGBS as follows. In the case of GBS, because it is so complex and the evaluation is many-layered, it would be inappropriate to think in terms of a single overall counterfactual. Rather, it is appropriate to consider what is the **appropriate counterfactual for each of the sub-enquiries** that make up the overall study. Following this approach to counterfactuals, the guidance on Evaluation Questions prepared for the country evaluation teams included the matrix below. However, the guidance also stresses that deciding what is the appropriate counterfactual, in principle, does not mean necessarily that it is practical to model one in econometric detail.

The two columns of the matrix show the possible counterfactuals to be considered for each main evaluation question (EQ). In practice, alternative aid modalities feature strongly among the counterfactuals.

EQ / LEVELS/ & DAC CRITERIA	POSSIBLE COUNTERFACTUALS
1. How does the evolving PGBS design respond to the specific conditions, strengths and weaknesses of the country, to government priorities and to the priorities and principles of the international partners?	Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.
2. Has PGBS contributed to greater harmonisation and alignment of the aid process?	What would have been the difference to H&A in the absence of PGBS (most likely alternative scenario)?
3. How efficient, effective and sustainable has been the contribution of PGBS to the performance of public expenditures?	Same aid volume, alternative modalities; or less aid.
4. How efficient, effective and sustainable has been the contribution of PGBS to improving government ownership, planning and management capacity, and accountability of the budgetary process?	Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.
5. How efficient, effective and sustainable has been the contribution of PGBS to improving public policy processes and policies?	Less aid. Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.
6. How efficient, effective and sustainable has been the contribution of PGBS to macroeconomic performance?	Same aid volume, alternative modalities, Less aid.
7. How efficient, effective and sustainable has been the contribution of PGBS to improving government performance in public service delivery?	Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.
8. How far has PGBS strengthened government impact on poverty?	Less aid. Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.
9. Is the PGBS process itself sustainable?	Structural adjustment type BS; other BS; other forms of aid; alternative designs of BS; alternative combinations of PGBS with other forms of aid.

For each main Evaluation Question in the country reports, a final section summarises the relevant counterfactual considerations.

The country reports' reflections on counterfactuals are very guarded and reserved because:

- the appropriate counterfactual is a matter of conjecture, rather than an alternative scenario that can be reconstructed in detail;
- a complete substitution of one modality for another is rarely seen as a realistic option; and in any case -
- many of the more interesting observations are about the interactions between different aid modalities.

Explicit consideration of counterfactuals thus encouraged reflection on complementarities as well as possible substitution among aid modalities.

* IDD and Associates (2007). *Joint Evaluation of General Budget Support 1994–2004 - Note on Approach and Methods*.

Practice, approaches and methods used by Multilateral Development Banks (MDBs) for measuring impacts of Policy-Based Lending (PBL) operations*

Among the most significant approaches to addressing PBL impact assessment and to discussion of the counterfactual, the following cases are mentioned:

- ✓ **AfDB Approach to assessing impacts of PBL reforms: the ‘before’ and ‘after’ method.** The most straightforward approach in evaluating PBLs is the "before-and-after" method that involves comparing base values of policy targets with values achieved at completion and post-evaluation. One application of this method is to measure indicators of the key development objectives and goals of the operation over a period of 3-4 years *before* the start of the operation, and for a similar period beginning near or at the end of the operation. This approach however, presents a key limitation: by itself it cannot separate out the effects of other factors (e.g. shocks, macroeconomic developments, other policy reforms, and changing market conditions can influence the same set of performance variables as PBL). Such an approach can be effectively used when the counterfactual (i.e., what would have happened without reform) is that the policy context remain the same, response lags are relatively short (i.e., less than 2 years after completion), and exogenous conditions prior to and after reform are approximately the same. The problem, however, is that the assumptions underpinning the validity of these comparisons may not hold.
- ✓ **Isolating the effects of PBLs from shocks through quantitative techniques.** Sophisticated economic analysis tools may be needed to untangle the direct and indirect effects that a PBL operation has on economic performance. Various quantitative modelling approaches can be used, ranging from partial equilibrium analysis to representative household models, simulation models, and applied general equilibrium models. The main advantage of employing such models in the evaluation process is their theoretical rigour. The main disadvantages are that they are demanding in terms of data requirements, time, and skills necessary for the evaluation. While the use of model-based evaluations is common in research-oriented exercises, in only a small number of cases have mandatory MDB evaluations employed formal economic models to assess PBL outcomes and impacts. Some PBL evaluations have employed simple growth decomposition methods to isolate the effects of policy change from major shocks and changes in the terms of trade. Others have used cross-country regression models to distinguish the effects of policy change from those related to the starting conditions and structural characteristics of borrowers.
- ✓ **Isolating the effects of PBLs from shocks through qualitative techniques.** Qualitative techniques can also be used to distinguish the effects of a shock from the effects of PBL. In a qualitative approach, evaluators draw on local knowledge and expertise to reach a judgment about the significance of the shock on the key outputs, outcomes, and impacts that a PBL operation aimed to achieve. A judgment is also reached on the degree to which those key results are likely to be achieved in the future, once the effects of the shock have passed.
- ✓ **Comparing to the Counterfactual Case.** A comparison with what would have resulted had there been no PBL operation —to a counterfactual— is the appropriate way to evaluate outcomes and impacts. This is a more analytically demanding and contentious task than “before-and-after” comparisons. The main problem is that the counterfactual is unobservable, and different stakeholders may have diverging views on what would have happened in the absence of any operation. Serious construction of counterfactuals is relatively rare in MDB evaluations. There are cases, however, in which a consensus of informed experts can be used to define a suitable counterfactual. In some cases, PBL appraisal reports have included a “without reform” scenario for certain key outcome and impact variables. It may also be possible to use performance of policy and impact variables in “similar” countries that did not undertake PBL-supported reform as a baseline comparator.

A simple comparison between what was anticipated and what actually transpired is sometimes sufficient to illustrate the order of magnitude of the likely PBL impacts. For example, in cases in which the counterfactual was an unsustainable situation (e.g., inflation out of control, imminent debt default, systemic collapse of the banking system), a comparison of the main performance indicators with the “untenable” counterfactual is sufficient to assess the effectiveness of a PBL operation.

* Tabor, S. R. (2005) *Good Practices for the Evaluation of Policy-Based Lending by Multilateral Development Banks*?. Report prepared for the Evaluation Cooperation Group of the Multilateral Development Banks.

Assessing Budget Support with Statistical Impact Evaluation

Elders, C., Gunning, J.W., and de Hoop, K., (2007)

In this paper the authors argue that existing statistical impact evaluation techniques can be modified in such a way that they become suitable for sector or general budget support evaluations.

The proposed methodology – applied by the authors to the education sector - envisages the following steps:

- Drawing of a representative sample of the target population (e.g. drawing of a sample of schools making the probability of being included in the sample proportional to the population size of a school's catchment area);
- Collection of intervention histories for the sample schools (record of policy-induced changes at the school level (e.g. availability of textbooks, number of class rooms, level of training of teachers / heads of the school, facilities, etc.));
- Collection of data histories on results (e.g. exam scores of the pupils of the sample schools);
- Identification of variables which may influence these results (other than the policy variables);
- Regression of changes in exam scores on changes in the intervention history variables.

The regression results can then be used to obtain an estimate of the aggregate impact of all the various schooling interventions.

The method provides an ex post assessment, ie. it addresses the question 'whether the money spent on education (for example) in a particular period was well spent' in the sense that it achieved a significant and substantial improvement in terms of exam results. The evaluation would further inform not only on the average return on educational spending but also on whether the portfolio of activities within the sector is efficient as it would reveal differences in returns between various government activities. The same reasoning could also be applied to infer differences in returns across (rather than within) sectors thus providing grounds for possible changes in the allocation of resources across activities / sectors to increase aggregate return.

This method – by providing an assessment of the effectiveness of a set of interventions undertaken within a particular sector and period - allows to infer that if donors have contributed to the budget of a given ministry it would be sensible to attribute the effect of those interventions to donors in proportion of their contribution.

It is however important to note that whereas this method has achieved significant results in assessing the relative importance of individual policy variables, it is unable to catch the impact of systemic reforms, which are often the consequence of budget support at macro and/ or sectoral level.